DOCUMENT RESUME

ED 117 200 TM 005 061

AUTHOR Hall, James W.; And Others
TITLE Further Development and Refinement of DIAL; State of Illinois Contract L171. Final Report.
INSTITUTION Illinois State Office of the Superintendent of Public Instruction, Springfield. Dept. for Exceptional Children.
NOTE 102p.; Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.76 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Data Collection; Diagnostic Tests; Intervention; *Predictive Validity; Preschool Education; *Preschool Tests; School Readiness Tests; Scoring; *Screening Tests; Test Construction; Testing; *Test Reliability
IDENTIFIERS *Developmental Indicators Assessment Learning; DIAL

ABSTRACT

In 1970, the Illinois Legislature passed bills requiring that elementary school districts institute procedures to identify and treat preschool children who, without such intervention, would not progress satisfactorily in a normal school environment. Efforts in this regard resulted in the screening instrument called Developmental Indicators for the Assessment of Learning (DIAL). Despite initial administrations of the battery and several research efforts, information regarding the reliability and predictive validity of the DIAL was insufficient to promote its general use for preschool screening. This project was funded in an effort to supply the missing data and to address additional questions. The major activities and findings of the project are described in the second section of the report. These concern the content validity of the DIAL as judged by consulting experts in child behavior and development, the predictive validity of the DIAL, the reliability of DIAL administration and scoring procedures, and various recommended modifications of the DIAL. The final section of the report summarizes conclusions and outlines recommendations concerning subsequent steps in the development and use of the DIAL. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED117200

FINAL REPORT

FURTHER DEVELOPMENT AND REFINEMENT

OF DIAL

STATE OF ILLINOIS CONTRACT L171*

OFFICE OF THE SUPERINTENDENT OF PUBLIC INSTRUCTION

DIVISION OF ACADEMIC AFFAIRS

DEPARTMENT FOR EXCEPTIONAL CHILDREN

SPRINGFIELD, ILLINOIS 62706

BEST COPY AVAILABLE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

James V. Hall

Carol D. Mardell

John W. Wick

Dorothea Goldenberg

*There are several auxiliary DIAL studies either just completed or still underway whose results could not be included in this report. Dr. Sam Mikaelian, Assistant Superintendent for Pupil Services, Wilmette Public Schools, Wilmette, Illinois, has assembled a cost analysis for the training and administration of DIAL. He is currently preparing a comparison cost structure between an early remedial programming budget and special education placement costs. Both of these reports should enable school personnel to evaluate the critical costs of an early screening program.

Mr. Lew Souff, psychologist and Walden doctoral student has recently completed his dissertation titled "A Comparison of DIAL Variables to Chronological Age and the Peabody Picture Vocabulary Test." That study may serve as a guide for continued longitudinal study or as an impetus for other directions in early childhood research.

Also, included as addenda to this report are the Project Summative Evaluation Reports, a report of the field use of the DIAL in Northbrook and Wilmette, and a copy of the DIAL Manual.

TM005 061

Table of Contents

Acknowledgements

I. Background and Introduction

II. Project Activities and Findings

A. Site Visits and Consultations with Specialists

B. Predictive Validity

C. Accuracy of DIAL Administration and Scoring

D. Recommended Modifications of the DIAL

III. Conclusions and Recommendations

Addenda

Concurrent Diagnostic Evaluation of DIAL II Children (Northbrook/Wilmette)

Summative Evaluation Reports (Lalor & McGrady)

DIAL Manual (Included only in original copy of report)

ACKNOWLEDGEMENTS

The present contract period was filled with the difficulties of collection and evaluation of a very large amount of information. Children had to be tracked down and data gathered. Computer cards had to be punched and repunched before final analyses could be conducted. None of this could have been accomplished without a great deal of cooperation among all of the parties associated with the project.

The project staff, Dr. James W. Hall, Project Director, Dr. Carol D. Mardell, Associate Director, Dr. John Wick, Data Analyst, and Ms. Dorothea Goldenberg, Research Associate, wish at this time to thank the following persons:

Formative Evaluator:
Dr. John Bormuth, University of Chicago

Summative Evaluators:
Dr. Ida Lalor, School District 65, Cook County
Dr. Harold McGrady, Director of Special Education, Mesa, Arizona

Consultants:
Dr. B. Bannister, Northwestern University
Dr. Bettye Caldwell, University of Arkansas
Dr. D. Clapper, Northwestern University
Dr. Leonard Feldt, University of Iowa
Dr. Lolas Halverson, University of Wisconsin
Dr. Doris Johnson, Northwestern University
Dr. T. E. Jordan, University of Missouri
Dr. Burton White, Harvard University

Staff:
Ms. Joanne Benge
Ms. Cheryl Gay
Mr. John Nordstrom

I. BACKGROUND AND INTRODUCTION

The major purposes of the present project were to provide data regarding the validity and utility of the DIAL and recommendations for the future development and use of that battery. Our efforts in these directions can be better understood by first sketching the background and context in which they occurred.

In 1970 the Illinois Legislature passed bills (HB322 and HB323) requiring that elementary school districts institute procedures to identify and "treat" preschool children who, without such intervention, would not progress satisfactorily in a normal school environment. The Office of the Superintendent of Public Instruction (OSPI) in Illinois undertook to assist the schools in complying with that requirement by an attempt to identify available procedures by which schools could screen preschoolers as an initial step in the required process. Specific responsibility for the discovery of appropriate screening procedures was assigned to Carol Mardell and Dorothea Goldenberg, then employed by OSPI and this past year members of the current Project Staff. Following a search of available material Dr. Mardell and Ms. Goldenberg concluded that adequate screening procedures were not currently available, and they undertook to develop such procedures. Their efforts resulted in the initial version of the screening instrument called Developmental Indicators for the Assessment of Learning or DIAL. To distinguish between that battery and the present version of it, the labels DIAL I and DIAL II will be used in this report.

In the spring of 1972 the DIAL was administered to some 5,000 children from 2.5 to 5.5 years of age. Subsequently, and based largely on the results of that extensive application, the DIAL was revised substantially and administered in the spring of 1973 to a second large sample of preschoolers in Illinois. To facilitate communication we shall refer to the original DIAL and the 1973 revision as DIAL I and DIAL II respectively.

The first major report concerning the DIAL was published in 1972.[1] That report, coauthored by C. Mardell and D. Goldenberg, described DIAL I and the procedures and rationale for its development. Also included were several reports of evaluators, which provided part of the basis for the revisions that resulted in DIAL II.

The second report on the DIAL[2] was authored by John W. Wick and published in 1973. That report summarized a substantial amount of information concerning DIAL II, and included a number of recommendations concerning the utility of the battery and desirable modifications of it. Also included were normative data and data concerning the stability over one year of DIAL scores.

Despite the efforts referred to above, information concerning the reliability and predictive validity of the DIAL was insuf-

---

[1]Learning Disabilities/Early Childhood Research Project, Annual Report to the Office of the Superintendent of Public Instruction, State of Illinois, August 31, 1972.

[2]Wick, John W. et al. Validation and Normative Study of the D.I.A.L. Battery. Final Project Report to The Office of the Superintendent of Public Instruction, State of Illinois, August 31, 1973.

ficient to promote its general use for preschool screening. Moreover, a number of questions concerning the content, administration, and scoring of the DIAL had been raised both by external evaluators and by individuals within the OSPI. The current project was funded in an effort to supply the missing data concerning reliability and predictive validity, and to address those additional questions.

Before moving to a description of Project activities and findings, we wish to make two points relevant to an understanding of this year's efforts. Both were made earlier in our interim report to the OSPI as well as in various meetings with OSPI officials concerned with the Project. First, we found during the initial weeks of the contract period an absence of sound long-range plans for the development and implementation of the DIAL. Guidelines for a number of specific decisions concerning the present Project simply were not available. As a result during the first three to four months a relatively large portion of the time of the Project Director and Formative Evaluator was devoted to searching for such guidelines, then, in their absence, to defining basic questions and issues that required resolution if optimal progress was to be made. The problem of broad long-range planning is one that remains, in our judgement, and further comments and recommendations concerning it are made in the final section of this report.

A second point concerns the specific objective of obtaining data concerning the predictive validity of the DIAL. The objective in examining the predictive validity of the DIAL battery was to

determine the relation between preschoolers' performance on that battery and their subsequent success in school, with particular attention to children who are unsuccessful in school. A fully adequate study would require that children tested with the DIAL battery be followed up at least as far as Grade 2, that the sample followed be varied with respect to factors (e.g., age, sex, SES) that may be related to the battery's validity, that the sample be uncontaminated in the sense that no intervention has resulted from the DIAL testing, and that adequate criteria of school success be defined and measured. In the present case only the last of these four requirements could be met. None of the children tested with the present version of the DIAL battery were beyond the kindergarten level. Many of the children tested in 1972 with the earlier version of the battery were in Grade 1, but none were beyond that. The samples followed up were sufficiently variable in several respects (e.g., sex, SES), but not in one important respect--age. That is, children below 54 months of age when tested in 1973 were not yet in kindergarten, and those below 54 months when they were tested in 1972 with DIAL I are not yet in Grade 1. Any validity data thus would be restricted mainly to the upper half of the age range for which the DIAL battery was designed. Finally, feedback regarding DIAL performance (e.g., schools) so that the possibility of intervention resulting from DIAL testing exists. Moreover, such feedback also was provided in the case of the low-performing children in the 1972 sample.

It is largely because of the above constraints that we concluded that no <u>definitive</u> evaluation of the utility of the DIAL

battery was possible at the end of this contract period. Nevertheless, validity data collected this year could at least give a rough estimate of the relationships in question. Therefore, such a study was completed and is described in a later section of this report.

In the next section of the report, the major activities and findings of the Project are described. These concern the content validity of the DIAL as judged by consulting experts in various specialties concerning child behavior and development, the predictive validity of the DIAL, the reliability of DIAL administration and scoring procedures, and various recommended modifications of the DIAL. The third and final section of the main body of the report summarizes our conclusions and outlines our recommendations concerning subsequent steps in the development and use of the DIAL. Several addenda also are included so that certain details and related material could be included without unduly increasing the length of the main report.

II. PROJECT ACTIVITIES AND FINDINGS

A. SITE VISITS AND CONSULTATIONS WITH SPECIALISTS

Field Visits

A list of potentially useful sites was made based on staff knowledge, interviews with various individuals, and a search of current related literature. We then generated and applied several criteria and procedures by which we determined the particular sites to be visited. Only sites that were actively engaged in attacking questions of early identification were considered as likely candidates for visitation. Second, we eliminated sites from which we could derive full benefits by obtaining publications or by long-range correspondence without a personal visit. Third, we attempted to avoid redundancy by eliminating sites that were likely to overlap heavily in terms of what they were likely to provide for us. The principle followed was to find out as much as we could via telephone calls, written correspondence and publications, and then to arrange visits only to those sites that met the criteria described above. Our purpose in each case was to glean from others information that would lead to improvements in the DIAL battery.

We anticipated making a minimum of five such visits, and perhaps as many as eight. However, these plans were altered following the first three visits made. Those visits, while interesting, did not seem sufficiently productive to justify additional expenditures of this sort. What we seemed to be learning, primarily, was that the DIAL Project was as advanced or more advanced in dealing with our problems than were others.

Visits by Consultants

The procedure used to identify and select consultants was similar to that described earlier for site visits. A long list of potential consultants was reduced to a handfull of well-known individuals representing somewhat differing backgrounds and perspectives regarding child development and behavior. We finally selected and contacted Dr. Bettye Caldwell, University of Arkansas, Dr. Leonard Feldt, University of Iowa, Dr. Lolas Halverson, University of Wisconsin, Dr. T. E. Jordan, University of Missouri, and Dr. Burton White, Harvard University. Consultation sessions were arranged and these meetings and the individuals involved are described below.

Consultation with Dr. Leonard Feldt, Professor of Education, University of Iowa on January 30 and 31 at O'Hare Airport in Chicago. Dr. Feldt is a widely-known expert in testing and measurement, and offered a number of recommendations regarding scoring, analyses, and the design and procedures for follow-up validity studies.

Consultation with Dr. Lolas Halverson, Professor and Director of the Motor Development & Child Study Center, University of Wisconsin on March 22, 1974, at Northwestern University. Dr. Halverson's area of specialization is motor development, and she inspected and evaluated those items in the gross and fine motor sections of the DIAL battery. Her

recommendations for revision centered mainly on administration and scoring procedures, and were not extensive.

Meeting with Dr. T. E. Jordan, Professor of Education, University of Missouri on February 6, 1974, at St. Louis, Missouri with D. Goldenberg. Professor Jordan has conducted an extensive longitudinal study of low-birth-weight children and made a number of suggestions concerning sample selection and assessment procedures in follow-up studies with children in the DIAL Project.

Meeting with Dr. Burton L. White, Professor of Education, Harvard University on March 8, 1974. Dr. White is well-known for his extensive research concerning development during infancy and early childhood and has developed a set of school success criteria in connection with his own longitudinal research program. He examined the DIAL battery in detail and offered a number of recommendations regarding the content of the battery and procedures for its empirical validation.

Meeting with Dr. Bettye M. Caldwell, Director, Center for Early Development and Education, University of Arkansas in Little Rock, on June 10, 1974, with

D. Goldenberg. Dr. Caldwell is well known for her research in early childhood education and as the author of the Cooperative Preschool Inventory. She made several suggestions which are noted in the manual revisions.

Additional Consultations and Site Visits

Relatively informal visits were made to two local sites. One of these is the District 65 Title III Early Childhood Program in Evanston which has been engaged in the development and use of a preschool screening instrument. The second was the Shaumburg Early Education Center in Shaumburg, Illinois where a program for preschool screening, diagnosis, and intervention has been developed. Arrangements were made for interchanges of information in the future. An informal consultation session concerning the motor sections of the DIAL battery was held with Professors D. Bannister and D. Clapper of Northwestern's Department of Physical Education. Several minor modifications of the gross motor sections were suggested.

Another evaluation was obtained from Dr. [illegible]is Johnson, Professor of Communicative Disorders, Northwestern University. Dr. Johnson had served on the advisory board of Project DIAL during its inception.

Summary of Expert Evaluation of DIAL Content

Of the two, the meetings with individual consultants have proven far more useful than have the site visits. Thus, we

did not conduct any further visits. We have obtained a number of recommendations that will be useful should the present battery be revised. In addition, the consultants employed and the project staff members of sites visited appeared to be in general agreement with the course of development of the DIAL battery. That is, the recommended revisions were, with few exceptions, minor in nature. One exception concerned the paucity of items concerning social-emotional development in the present battery. In the main, through, the revisions suggested were more similar to those that have been recommended on a number of other occasions (e.g., 1973 report by J. Wick) which seem likely to increase the reliability and validity of the battery without changing its character in any fundamental way. The minor revisions, along with a code indicating their source, are listed as part of the adjusted manual and can be implemented without further study. More extensive proposed revisions are included as well in a later section of this report.

II. PROJECT ACTIVITIES AND FINDINGS

B. PREDICTIVE VALIDITY

Sample Selection Considerations

The original DIAL I sample included 4,400 children, ages 30 to 66 months. This sample was obtained in the Spring of 1972 from throughout the state of Illinois. The sample selection goal was to obtain 96 kindergartners and 192 first graders. Certain constraints were employed in the selection, including the following:

1. The child must be currently enrolled in kindergarten or first grade.

2. The readiness test or achievement test, administered by the district, must be available for kindergarteners or first graders, respectively.

3. For data collection efficiency, at least nine other DIAL I screened children must be in the same school district.

4. Stratification design based on sex, race, urban/rural, and socio-economic status must be maintained with no more than two times as many children in any one cell compared to any other.

As the study developed, the possibility of administering the same achievement test to all children in the sample was considered. The distances between sites was large enough, however, to make this a very tedious and time-consuming task. Such a decision would have also required additional clearances--to get the child released from class and to convince parents that the testing was needed. After weighing the benefits of this approach against the apparent costs, the decision was made to use only

readily available data, as outlined by the four constraints listed above.

The problems of test bias since original DIAL I screening were considered minimal. For one thing, the DIAL I population's performances had not been previously scored, so that no formal results had been given to either parent or school staff member. The original score sheets of 1972 performance were not reviewed prior to sample selection in 1974.

Operational Definitions of the Categories

As noted before, the sample was stratified along four dimensions:

1. Race. In the DIAL I battery, a photo was taken of each child and attached to the score sheet. Based on this photo, each child was categorized as black or nonblack.

2. Sex. This information was reported on the score sheet.

3. Socioecoromic status. The categories here were labelled low and non-low. The children were separated into two socio-economic levels, using the 1960 United States Census Socioeconomic Status Scores for 297 occupational categories. Each target child's SES was evaluated numerically using for placement the father's current occupation as a first criterion and, if there was no record of this, the mother's occupation. Any occupation from 1 to 39 was considered low, while any occupation above 39 was considered non-low.

4. Demographic location. The categories here were urban and non-urban. Criteria for establishing demographic dichotomies were accomplished through use of the U.S. Census Bureau records.

Urban communities included populations above 50,000. Any community below this figure was considered to be non-urban.

As might be expected, the task of locating children two years after the time of the original testing was a formidable one. A target sample was first selected to complete the requirements of sample size with the various dichotomies described above. In cases where data were not available on a target child, that child was replaced from a pool of alternates. In some cases, too few alternates were available and certain cells were not completed. This was particularly true for the category black-non-urban.

Members of the project staff visited the schools where target children had been identified. After verifying the target child's attendance at the school, the teacher's measurement of certain behavioral indicators concerning the target child was obtained. An instrument (called the Classroom Performance Assessment) was given to the teacher for completion. The teacher was asked, in this instrument, to rate the target child on categories such as Disruptive, Emotional Control, and Inhibition. Twelve categories were included, in addition to an overall rating. The category names are listed in Table V of this chapter. A copy of the Classroom Performance Assessment is included in this chapter.

The teacher is the measuring device in this setting. Teachers differ in such ratings -- some are more harsh than others and will generally rate at a lower level. To provide a basis for comparing the relative rating of the target child by the teacher with others in the classroom, five cohorts were chosen at random from the teacher's gradebook. These five, plus the target child,

were listed on the C.P.A. in alphabetical order. The teacher was not told which of the six children was the designated one. The teacher thus provided information about the target child relative to ratings of a random sample of cohorts.

At the same time, school records of the child were examined to complete a data card. This card included information on achievement test results, date of testing, records of special education placement, diagnostic testing, and age. Teacher information recorded included the number of children in each classroom, number of years teaching experience, and sex of teacher. The latter was not a variable since all children in this study had female teachers.

The school personnel were given information about the goals and objectives of the project, how the children were selected, and the purposes of the longitudinal study. At no time did school staff see original DIAL I test data or scores.

The problems of time, distance and staff involvement necessitated that data be collected for alternates at the same time as regulars. The original design called for 96 kindergarteners and 192 first graders from 4 sites. Data were collected for 86 kindergarteners and 163 first graders. The reasons for the losses include 1) mobility of child and parent, 2) lack of identifying information, 3) time constraints, and 4) parent refusal. The actual sample obtained is listed in Table I.

Information about Data Collection at the Sites

Each site was contacted to determine if the original DIAL coordinator was still available. Follow-up efforts were initiated

Table I

Sample Actually Obtained
According to Site and Category

			NON-URBAN		URBAN	
			M	F	M	F
BLACK	LOW	K	2-D 3-J 5	3-J 1-Q 4	6-J 6	6-J 6
		1st	5-D 1-J 6	4-D 2-J 6	11-J 11	12-J 12
	NON-LOW	K	2-D 2-J 2-Q 6	1-D 2-J 1-Q 4	5-J 5	5-J 5
		1st	2-D 2-J 3-Q 7	3-D 1-J 1-Q 5	12-J 12	12-J 12
NON-BLACK	LOW	K	4-D 2-L 6	5-D 1-L 6	6-J 6	4-J 4
		1st	11-D 1-L 12	7-D 5-L 12	12-J 12	12-J 12
	NON-LOW	K	6-L 6	6-L 6	5-J 5	6-J 6
		1st	2-D 9-L 11	12-L 12	11-J 11	10-J 10

L - Lake County

D - DeKalb County, Elgin, Freeport

J - Joliet, Posen Robbins

Q - Quincy

based on recommendations at the site. Most of the sites were cooperative with the follow-up study of DIAL.* A range from irritation and resistive efforts to pleasant and meticulous pursuit of children's location was encountered. If the administrative structure was resistive because of internal difficulties, the path became encumbered. The great number of individual school districts, differing administrative structures, procedures, and regulations under which target children fell presented the most complicated communication problems. Once the goals and directions of the project were carefully explained to a local administrator, who had authority to expedite the communications, all further contacts moved fairly smoothly. Table II is a summary of population characteristics and Table III shows the schools contacted.

Sites

1. Lake County**

The first site selected was Lake County because of its close proximity to Northwestern University. The goal was a nonblack, nonurban population of 24 kindergarteners and 48 first graders, evenly divided by sex and SES characteristics. The original DIAL screening in Lake County was done in two schools. All children screened were of preschool age then. Some thirty-four school districts exist in Lake County -- not to mention various parochial

*The original contract with each site included a stipulation to provide data for five years; schools were obligated by law to participate.

**The full cooperation of Mr. Larry Vuillemot, director of Special Education District of Lake County (SEDOL) who assisted in the original DIAL testing as well, was obtained.

Table II

Population Characteristics
Used to Facilitate Sample Selection

SITE	POPULATION CHARACTERISTICS			# SOUGHT			# FOUND			% OF # SOUGHT	% OF TOTAL SAMPLE		
Joliet	Urban			144			135			94%	5[illegible]%		
		Black	Nonblack		72	72		69	66			28%	26%
		Low	Nonlow		[illegible]	72		69	66			28%	26%
		Kgn	First		48	[illegible]		43	92			17%	37%
Lake County	Non-			72			71			99%	29%		
	Urban	Black	Nonblack		0	72		0	71			0%	29%
		Low	Nonlow		36	36		36	35			15%	14%
		Kgn	First		24	48		24	47			10%	19%
Elgin	Non-			72			43			60%	17%		
Freeport	Urban	Black	Nonblack		72	0		43	0			17%	0%
Posen-Robbins		Low	Nonlow		36	36		21	22			8%	9%
Quincy		Kgn	First		24	48		19	24			7%	10%
				288			249			86%	100%		

Table III

Schools Attended By Sample

SITE	SCHOOL NAME
Lake County	Avon Hooper Plesiak Sprague
DeKalb	Big Rock Byron Genoa Hinckley Leaf River Paw Paw Rochelle Shabbona Somonauk Waterman
Elgin	Channing Hillcrest Huff
Freeport	Douglas Empire First Ward Henney Taylor Park
Posen-Robbins	Childs Turner
Joliet	Culbertson Farragut Forest Park Jefferson Keith Kelly Marsh Marshall McKinley Parks Taft Thompson Woodland
Quincy	Dewey Madison Washington

schools -- and the children could have been in any of those. A letter was sent to each school district superintendent asking for cooperation. Using addresses as a guideline, 12 school districts which might have a cluster of DIAL children were visited. Based on this investigation, three school districts with the largest clusters were identified. Data gathering was restricted then to District 41-Lake Villa, District 47-Avon, and District 103-Lincolnshire-Prairie View. Within these three school districts, data were collected on 42 children.

2. DeKalb

The second site selected was DeKalb because their original contract was to screen nonurban, low SES, black and nonblack children.* Due to the vast geographical area to be covered and the absence of coordination available between sub-sites, two teams set out to revisit communities where screening had been available during the original DIAL I data collection.

As a result, target children were located in the following communities: District 145-Freeport, District 46-Elgin, District 231-Rochelle, District 226-Byron, District 429-Hinckley Big Rock

---

*The initial site selection for obtaining an urban low and non-low sample of kindergarten and first grade children was Chicago. In the original DIAL design, Chicago furnished 1250 children and was in close proximity for follow-up and data collection.

Preliminary contact provided approval for Chicago school participation. However, the collection of data abruptly halted when it became apparent that complete data collection would be virtually impossible for both the kindergarten and first grade subjects. This was due to the lack of a school testing program in these grade levels. In many schools, there was no specific plan or design for achievement testing in first grade although readiness scores for kindergarten were available.

District 425-Shabbona, District 432-Somonauk, District 270-Leaf River, District 431-Waterman, District 271-Lee Center, and District 424-Genoa Kingston. Within these eleven school districts, data were collected on 48 children.

3. Joliet*

The entire sample urban population was obtained in District 86-Joliet. Within 13 schools, data were collected on 135 children. In addition, a cluster of nonurban, black children who had originally been part of the DIAL site in Joliet actually resided in District 143-Posen Robbins. Within this school district data were collected on 16 children.

4. Quincy

The total validation design required follow-up data on a nonurban black sample of children. This particular group of subjects was extremely difficult to re-locate. Within District 172-Quincy data were collected on 8 children.

The sample distribution by sites is as follows:

1. Lake	42
2. DeKalb/Elgin/Freeport	48
3. Joliet/Posen Robbins	151
4. Quincy	8
Total	249

*We appreciate the complete cooperation of Dr. Donald D'Amico, Superintendent; Mr. Martin Kinert, Director of Special Education; and Mr. Robert Stoner, Chief Psychologist.

Results

The first general use of the DIAL battery (DIAL I) occurred in the Spring of 1972. This screening led to the revision of many of the DIAL I items, and in the Spring of 1973 another substantial sample was screened with a revised instrument, DIAL II. Additionally, in the spring of 1973, some 500 children originally screened with DIAL I in 1972 were rescreened with DIAL II. In this section this group of 500 children will be called the 1973 Validity sample.

For a complete description of and rationale for the scoring system developed, the reader is directed to Section VI - Data Analysis in the "Six-month report of Project DIAL Activities under State of Illinois Contract No. L171" by J. Hall dated May 15, 1974, and to the final report by J. Wick, dated August 31, 1973, covering the validity study carried out under his direction.

Comparison of Scoring Systems

As is reviewed in the above documents, a considerable amount of effort has been expended in the development of the scoring system. Three systems are discussed:

1. A Raw Score system, in which the actual response on an item is treated as the score. This system has the difficulty of differentially weighting items as a function of number of response positions rather than as a function of the importance of the item.

2. The Score Sheet system. After the 1972 DIAL I data were gathered, a revised system was developed which weighted items on

a 0-3 or 0-2 scale. This system eliminated some of the problems of the Raw Score system, as item variabilities were diminished. Response groupings were based on the best data available at the time the revisions were made. Many times a single response could gain a score of 3, while other groupings which also allowed a maximum score of 3 contained as many as 12 or more single responses.

3. The Weighted system. Essentially, this weighted scoring system attempts to weight items objectively such that those items which best predict later learning problems are weighted most heavily.

In investigating the relative merits of different scoring systems, the primary criterion measure was the standardized test information gathered from the 249 children earlier in this chapter.

Table IV shows the relative predictive power of the two scoring systems. Twenty different criterion measures are used. These are described later. For the Score Sheet column, the four part-scores on the DIAL battery were used as independent variables, and the various test scores (from 1974) used as dependent measures. For the Weighted Scale column, the four weighted DIAL part-scores were used as predictors (independent measures) and the various test scores from 1974 were dependent measures.

A whole series of other comparisons of the weighted versus unweighted system were run. Correlations between weighted and unweighted DIAL II Scores and reading readiness data and correlations between the weighted/unweighted scores and teacher behavioral

Table IV

Comparison of Multiple Correlations with Various Criterion Measures Using Score Sheet and Weighted Scale Systems*

Criterion Measure	Score Sheet	Weighted Scale
1	.48	.53
2	.51	.53
3	.50	.54
4	.52	.58
5	.47	.55
6	.51	.52
7	.52	.53
8	.49	.56
9	.53	.58
10	.53	.59
11	.51	.53
12	.53	.60
13	.50	.56
14	.54	.54
15	.50	.53
16	.52	.57
17	.51	.57
18	.53	.57
19	.52	.55
20	.53	.59
Average	.51	.56

*The average multiple correlations for the various criteria is shown. In general, the weighted scale raises the multiple R by about 10%.

ratings of children were included in these "other" computations. In general, the results were consistent in all runs -- a 10% average increase in the correlations. The conclusion seems clear that some alteration to the current scoring system which appears on the score sheet will increase the predictive accuracy of DIAL.

Future Plans with Scoring System

As revealed in earlier reports, the weighting system is really a bit cumbersome. In the first place, the weights are a function of age and sex. Given these entrance parameters, each response by the child is assigned a weighted value for that item.

The weighting system which leads to the higher multiple correlations has one primary underpinning. Those items which are the better predictors should have the most impact on the prediction. This is accomplished by manipulating the amount of variability in the weighted scores. Those items with the most variability will have the greatest impact on the prediction. In addition, the weighting system essentially regrouped the items by treating each response as a separate item -- rather than the somewhat arbitrary groupings found on the score sheet.

Before the DIAL is used again, the score sheet should be revised to make use of the benefits of the weighting system while maintaining the relative simplicity of the unweighted approach. That is, the items will be regrouped, and the variability of possible scores manipulated, both on the basis of currently available empirical evidence.

The current score sheet system allows only three score points (0-2) or four points (0-3). No practical difficulty can

be forseen if this were changed to variations of two points (0-1) to six points (0-5). This will allow standard deviations of individual item scores to vary by as much as a factor of 5. The item regrouping and rescaling will be accomplished toward the goal of not diminished currently-attained multiple correlations.

Discussion of Criterion Measures Used in DIAL I Analysis

The procedures in the follow-up efforts on children originally tested with DIAL I are described earlier. A variety of standardized test scores were collected. These came primarily from four different batteries.* Even from within one battery (for example, the Iowa Test of Basic Skills) different subtests and different scaled scores were reported. The data were recorded in the field in this general format:

Subtest 1: A standard score
A grade equivalent
A percentile rank
A stanine.

This was repeated for three other subtests. In general, subtests 1 through 3 were reading-based subtests, and the math-based subtests were reserved for subtest 4 in the coding sheet. Because of the different tests administered, subtests chosen by the district, and score-reporting systems chosen by the district, the data available began something like this, where "bb" means no

*Iowa Test of Basic Skills, Metropolitan Achievement Test, Stanford Achievement Test, Metropolitan Reading Readiness Test.

information: (note that children 1, 2, and 3 may have had different batteries as well)

Child 1: bb xx bb bb bb bb bb bb bb xx bb bb bb xx bb bb

Child 2: bb bb xx xx bb bb xx xx bb bb bb bb bb bb xx xx

Child 3: bb bb bb xx bb bb bb xx bb bb bb xx bb bb bb xx

In order to facilitate the analysis (and not lose any of the data which were collected) a series of assumptions and estimates were used as criterion scores. These were:

1. The publisher's battery was used to convert into the other scores for a given battery. Thus, in the case of Child 1, the grade equivalent available was used to fill in the other scores.

2. For percentiles to stanines (and vice versa) commonly accepted conversions were used.

3. Where a subtest was missing, the average of standard scores available for that child was inserted.

4. The standard scores were all converted to a common score.

5. For the grade equivalents, percentile ranks, and stanines, the assumption was made that all systems were based on essentially the same normative population.

All of these assumptions -- and particularly the fourth one -- are likely to introduce error variance into the multiple regression estimates. Stated another way, it is highly probable that the multiple correlations given are lower than would have been found had the same battery been used for all of the children. It is difficult to imagine how these asuumptions could have enhanced the multiple R's in any way.

In any event, the criterion measures from Table IV can now be explained. Criteria 1 through 4 are the four scores on subtest 1 (a standard score, grade equivalent, percentile rank, and stanine), 5 through 8, 9 through 12, and 13 through 16 the four scores for subtests 2, 3, and four, respectively, and 17 through 20 the means for the four subtests. You will note that we have marked an arrow by the five different stanine scores. None of the scaled-score systems seems dominantly superior to the others, and for purposes of simplicity we have decided to report only stanine scores as the dependent variable for the remainder of this section.

Prediction to Subgroups Within the Total Sample

Data were gathered such that each child could be placed in one of two mutually exclusive categories along four dimensions. The manner in which the categories were defined is found elsewhere in the report. The four dichotomies and the multiple correlations to each, are found in Table V below:

Table V

Breakdown of Dichotomies - Multiple R's*

Def.	Sub.1	Sub.2	Sub.3	Sub.4	Avg.	Max.
Male	.59	.57	.60	.58	.60	.60
Female	.51	.58	.54	.53	.53	.54
Black	.58	.54	.58	.59	.59	.59
Nonblack	.51	.52	.54	.50	.53	.54
Urban	.56	.62	.64	.61	.63	.64
Nonurban	.55	.52	.54	.53	.55	.55
Lo SES	.72	.70	.72	.73	.73	.73
Non-lo SES	.58	.45	.50	.47	.49	.50

*dependent is subtest stanine, four independent are weighted DIAL subtest scores.

A comparison of Table IV in which the entire follow-up sample is treated as a unit with Table V seems to suggest, at least, the following:

1. In general, breaking the total groups into the four dichotomies does not dramatically change the multiple correlations.

2. Lower original DIAL I scores lead to higher multiple correlations. Previous analyses have shown that the four first named categories in the dichotomies (i.e., males, blacks, urban, and lo SES) scored lower on the original DIAL screening.

Analyses on further subdivisions were run (for example, male-black or lo-SES-urban. As the original sample (N approximately 250) is subdivided, the sample sizes become more and more variable, and the correlations similarly become more erratic. The results from further subdivision tended to be in the range shown by Tables IV and V.

One is led to the tentative conclusion that the multiple correlations are more sensitive to range of scores than to any of the particular subgroups defined by the dichotomies. To test this conclusion, we determined the range of initial weighted DIAL I scores for each of the dichotomies, and ranked them from 1 to 8 (1 = smallest range, 8 = largest range). The same ranking was applied to the multiple R's. Here is a table of the results:

Table VI

Observed Maximum Multiple Correlations for the Eight Subgroups and Relative Range of Initial DIAL I Scores

Def.	Rank of Multiple R	(Multiple Correlation)	Range of Scores (rank)
Non-lo-SES	1	(.50)	1
Non-black	2-3	(.54)	2
Female	2-3	(.54)	4
Non-urban	4	(.55)	3
Black	5	(.59)	6
Male	6	(.60)	5
Urban	7	(.64)	7
Lo-SES	8	(.73)	8

As can be seen, the correspondence is fairly clear. The size of the multiple correlation changes with the range of original scores. The only inconsistencies which are seen occur where the maximum multiple correlations are very close anyway. One additional piece of information is germane at this point: The range limitations are primarily caused by an absence of low scores in the groups with smallest multiple correlations.

The results to this point support the notion of a single set of normative tables rather than group-specific norms. In past reports, we have suggested the need for sex-specific norms at least. The 1972 item information showed quite clearly that on the majority of items the males were developmentally behind the females, a result which was interpreted as suggesting the need for sex-specific norms. However, public schools do not have -- or allow -- de jure separation by sex, location, race, or SES. Students go to school together, operating under the same general

environmental conditions. The definition of "failure to learn" -- which is what this test is attempting to predict -- is probably more a function of the environment of the classroom than the sex of the student within the individual classroom.

All of this seems to suggest a single set of norms, sensitive only to age. The definition of "high risk" would be left to the individual system, where some sort of flexible "quota system" would be established -- say 20% of some areas of environmental deprivation versus 5% in more affluent areas. Note that this is a very tentative suggestion. Data collected in the next year -- when the DIAL I sample moves to second grade and the DIAL II group reaches first grade -- should lead to some more objective decisions along these lines.

DIAL II Scores as a Predictor of Reading Readiness Scores

One group of 193 children who had taken DIAL II in 1973 provided data on performance on the Metropolitan Reading Readiness Test. The sample was taken entirely from a northern suburb of Chicago. Ages, at time of testing, ranged from 52 to 65 months. Only the four subtest scores were available for the analysis -- we have called these the "score sheet" system or unweighted scores. These four scores were used as independent variables. The criterion measure was the percentile rank of the child on the Metropolitan Reading Readiness Test. The results are as follows:

Table VII

Results Using Unweighted DIAL II Scores as Predictors and Reading Readiness Percentile Ranks as Criterion

Independent Variable	Entered	Multiple R	Simple R
Gross Motor	1st	.432	.432
Communications	2nd	.532	.462
Fine Motor	3rd	.598	.550
Concepts	4th	.640	.540

Another run was made which may be of interest. The age of the student (in months) was added as a fifth independent variable. Although the simple correlation between age and MRRT scores was 0.40, the multiple R was not increased at all. For this sample, the four DIAL II part scores account for the variance without taking age into consideration as well as when age is considered.

The results are probably low for two reasons. First, as was explained earlier, the weighted DIAL scores tend to lead to multiple correlations about 10% higher. Additionally, this sample was taken from a fairly homogeneous population. We can expect that the range of DIAL II scores obtained is restricted -- probably at the lower end. A more complete distribution of scores across the scale continuum probably would also have led to higher multiple correlations. The "true" multiple correlation is more likely to be somewhere between 0.70 and 0.80.

Correlations Between DIAL I Scores and Teacher Ratings

In the follow-up studies of DIAL I children who were in first grade during the 1973-4 school year, teacher behavioral ratings of students were obtained. The rating form, and the techniques used to obtain the information, are described elsewhere. Table VIII shows the results.

Table VIII

Multiple and Simple Correlations Between
Teacher Behavioral Ratings (Criteria) and DIAL I Scores

Criterion	Multiple R	Simple Correlations GrossM	FineM	Concepts	Comm.
Disruptive	.24	.00	.18	.02	-.02
Motor act	.29	.08	.26	.09	.03
Emotional control	.19	.05	.17	.07	.04
Coordination	.20	.06	.18	.05	.04
Group participation	.19	.12	.19	.11	.13
Perseverance	.23	.03	.18	.04	-.03
Compliance	.18	.05	.16	.01	.06
Attention	.21	.08	.19	.09	.03
Frustration level	.22	.05	.19	.01	.06
Peer relationships	.22	.07	.18	.00	.05
Inhibition	.22	.05	.18	.01	.05
Aggression	.22	.03	.19	.02	.06
Overall rating	.27	.15	.26	.10	.19

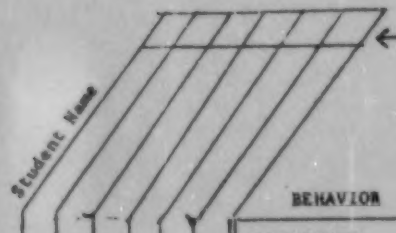
Two fairly obvious interpretations of Table VIII are these: The multiple correlations for the ratings are much lower than the correlations with the standardized tests, and the fine motor section of DIAL is clearly the best predictor of rating scores.

The low correlations are too suprising. DIAL I has no behavioral measures and the teacher interview form does not ask how the child will perform on learning tasks. The predictions are to two criteria which apparently are quite different.

DIRECTIONS: Please rate each child listed on the left for each behavior. Write the appropriate number (1-5) which best describes each child. Then give each child an overall rating (1-5) for your estimate of potential school success with 5 being the highest and 1 being the lowest.

CLASSROOM PERFORMANCE ASSESSMENT (C.P.A.)

Student Name / BEHAVIOR						
	DISRUPTIVE	Consistently out of seat which bothers or distracts others; consistently talks out without permission 1	Usually out of seat which distracts others; usually talks out without permission 2	Sometimes out of seat but not disturbing; occasionally talks out without permission 3	Seldom out of seat or talks out without permission 4	Out of seat only after asking and rarely talks out without permission 5
	MOTORIC ACTIVITY	Consistently fidgety or antsy - always in some kind of motion which distracts others 1	Usually overactive: i.e. tapping fingers, swinging feet, scratching, rubbing which can be distracting 2	Sometimes fidgety but just as often can inhibit motion 3	Seldom fidgety or antsy - able to control or inhibit motion 4	Highly controlled motoric activity; can sit still for long periods of time 5
	EMOTIONAL CONTROL	Reacts violently with no provocation 1	Usually uncontrolled emotions to slight provocation 2	May attempt control, but just as often will display feelings 3	Usually controls emotional outbursts 4	Highly controlled; chooses socially acceptable alternate behavior 5
	COORDINATION	Consistently awkward or worksheets are hopelessly careless 1	Usually clumsy; trips but does not fall; or work is usually slipshod 2	Sometimes clumsy; or inconsistent performance between careless and carefully done work 3	Usually produces careful and exact work; rarely clumsy in movements 4	Well coordinated; produces very exact and careful work 5
	GROUP PARTICIPATION	Never takes part in discussions 1	Rarely takes part in discussions 2	Occasionally takes part in discussions 3	Usually takes part in discussions 4	Consistently takes part in group discussions while still allowing others to participate 5
	PERSEVERATION	Always repeats what he is saying or doing when it is no longer appropriate 1	Usually repeats what he is saying or doing when it is no longer appropriate 2	Sometimes repeats what he is saying or doing when it is no longer appropriate 3	Rarely repeats what he is saying or doing when it is no longer appropriate 4	Never repeats what he is saying or doing when it is no longer appropriate 5



CLASSROOM PERFORMANCE ASSESSMENT (C.P.A.)

BEHAVIOR					
COMPLIANCE	Never does what another person has requested 1	Rarely does what another person has requested 2	Sometimes does what another person has requested 3	Usually does what is requested but there is a time delay 4	Always does what is requested as soon as it is asked 5
ATTENTION	Loses concentration at slightest distraction 1	Finds it very difficult to maintain concentration 2	Can concentrate for short periods 3	Good concentration; absorbed in one subject 4	Excellent concentration; can even focus on two things at once 5
FRUSTRATION LEVEL	Refuses to attempt tasks or says "I can't" without trying 1	Withdraws from situation when immediate success is not achieved 2	Withdraws from situation after second attempt is unsuccessful 3	Makes several attempts at activity before giving up 4	Copes well with extended frustrating situation 5
PEER RELATIONSHIPS	Avoided by all other students; unable to make friends 1	Avoided by most; rarely makes friends or accepted by subgroup with similar social problems 2	Accepted by most; occasionally makes friends 3	Liked by most; usually makes friends 4	Liked by all students; always makes friends 5
INHIBITION	Always acts (verbally & nonverbally) without forethought; does not consider consequences 1	Usually acts (verbally and nonverbally) without forethought; usually does not consider consequences 2	Sometimes acts (verbally and nonverbally) without forethought; sometimes does not consider consequences 3	Usually acts (verbally & nonverbally) with forethought; usually considers consequences 4	Always acts (verbally & nonverbally) with forethought; considers consequences 5
AGGRESSION	Consistently attacks or attempts to attack another person; consistently destructive to objects at school 1	Often attacks or attempts to harm another person; often destructive to objects at school 2	Sometimes attacks or attempts to harm another person; sometimes destructive to objects at school 3	Seldom attacks or attempts to harm another person; seldom destructive to objects at school 4	Never attacks or attempts to harm another person; never destructive to objects at school 5

II. PROJECT ACTIVITIES AND FINDINGS

C. ACCURACY OF DIAL ADMINISTRATION AND SCORING

The accuracy with which operators administered the DIAL and the reliability of scoring were examined using operators trained with procedures developed previously. Those training procedures were developed and applied by C. Mardell and D. Goldenberg with considerable attention to recommendations generated by the DIAL I training evaluation. Verbal information objectives and performance objectives were developed and a level of 80 per cent correct on both a written and a performance test was established as the criterion for successful completion of training. The training was conducted at a screening site over a two-day period with small groups of trainees.

During the first of the two days test materials, manuals, and the testing rationale were introduced. Operators were paired for role playing and further examination of the testing procedures as well as for practice presentation of individual station tasks. On the second day critical observations were made during the actual testing of children. Each operator was evaluated on the accuracy of individual task presentation and scoring. Individual feedback conferences then were held with each trainee.

DIAL Administration

The study described in the following paragraphs was conducted to provide more systematic data than previously available concerning inter-operator reliability of DIAL administration.

Method. The subjects were 4 white females trained as described earlier. The screening site was a middle class, mid-

western, suburban school district. Each testor was randomly selected from a core of trained personnel. A total of 14 children were tested by these operators. Eight children were male and 6 children were female. Among the children were seven 5-year olds, four 4-year olds, and three 3-year olds.

A tape recorder was placed at each DIAL station for collection of all verbal responses during test administration. Each testor was responsible for only one DIAL station. The recordings included morning and afternoon screening sessions. Every operator administered all items of the specific DIAL station to children randomly assigned.

Findings and Conclusions. The tapes were transcribed and the results tabulated as shown in Table 1 . As may be seen in that table, errors were relatively infrequent and generally minor. In most cases, the changes involved additions rather than omissions or substitutions. More reinforcing comments were used by the testor to encourage a child's response. Questions were recorded by the testors when the child's voice quality was inaudible or jumbled.

Every child's length of testing time was recorded in minutes and seconds. In most cases, the younger child took greater time to complete the DIAL items than did older children.

The four testors were consistent in use of the manual. Every testor successfully followed the manual according to entrance or exit items. There were no errors in establishing rapport. Every testor succeeded in providing a warm, verbal testing climate. There were some word changes or additions in each DIAL station

as seen in the figure.

In the gross motor area, Operator A had one word change during the screening of four children. There were no word changes in the fine motor by Operator B while screening two children. Operator C and Operator B scored on the same single error in the concepts area with a word substitution of "give" for "hand". During the screening in the concepts area many additions occurred. These additions were positive reinforcements or the clarification of a direction. In the communications area Operator D made the same error twice with two of the four children screened. There was a word substitution of the word "if" for the word "when".

The data just presented indicate a very high level of accuracy in the administration of the DIAL under the conditions of this particular study. These conditions were ones which would be expected to yield maximum performance, since operators were trained by the individuals most experienced and knowledgable concerning the DIAL (C. Mardell and D. Goldenberg) and since the operators were aware that their performance was being studied. It remains to determine the degree of accuracy of administration that can be expected under more typical field testing and how that accuracy would be affected by variation in training techniques. In fact, one of the higher priority tasks in the future development of the DIAL should be the development of feasible and effective training procedures.

Table 1

Summary of Accuracy of DIAL Administrative Procedures - Gross Motor

Operator	Child #	Sex	Time of Day	Provides Warm Climate	Uses Exact Wording	Word Additions Or Changes	Enters & Exits According to Manual	Child Age	Area Subtests
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Throwing
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Catching
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Jumping
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	-	on 1 ft. till I tell you to stop	+	5-1	Hopping
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Skipping
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Standing Still
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Body Parts
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Balance Beam
A	1	M	PM	+	+		+	3-7	
A	2	M	PM	+	+		+	5-1	
A	3	F	AM	+	+		+	4-5	
A	4	M	AM	+	+		+	5-1	Growth

Child #	Testing Time	Child #	Testing Time
1	7'20"	3	6'05"
2	5'10"	4	6'00"

Table 1

Summary of Accuracy of DIAL Administrative Procedures - Fine Motor

Operator	Child #	Sex	Time of Day	Provides Warm Climate	Uses Exact Wording	Word Additions Or Changes	Enters & Exits According to Manual	Child Age	Area Subtests
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+		+	3-3	Designs
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+		+	3-3	Blocks
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+		+	3-3	Cutting
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+	an you try	+	3-3	Drawing Shapes
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+	Show me—you try	+	3-3	Drawing Letters
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+		+	3-3	Finger Agility
B	9	F	PM	+	+		+	5-1	
B	11	F	AM	+	+		+	3-3	Hand Clapping

Child #	Testing Time
9	7'20"
11	7'25"

Table 1

Summary of Accuracy of DIAL Administrative Procedures - Concepts

Operator	Child #	Sex	Time of Day	Provides Warm Climate	Uses Exact Wording	Word Additions Or Changes	Enters & Exits According to Manual	Child Age	Area Subtests
B	5	F	AM	+	+		+	4-6	
B	7	F	PM	+	+	keep looking-I knew you'd find it	+	5-0	
C	6	F	AM	+	+		+	5-3	
C	8	F	PM	+	+	that was just wonderful	+	5-0	
C	9	F	PM	+	+	that's just perfect	+	5-1	
C	4	M	AM	+	+		+	5-1	Learning
B	5	F	AM	+	+		+	4-6	
B	7	F	PM	+	+	terrific	+	5-0	
C	6	F	AM	+	+		+	5-3	
C	8	F	PM	+	+	that's wonderful	+	5-0	
C	9	F	PM	+	+	perfect	+	5-1	
C	4	M	AM	+	+	very nice	+	5-1	Colors
B	5	F	AM	+	+		+	4-6	
B	7	F	PM	+	+		+	5-0	
C	6	F	AM	+	+	that's perfect	+	5-3	
C	8	F	PM	+	+	that's marvelous-just perfect	+	5-0	
C	9	F	PM	+	+		+	5-1	
C	4	M	AM	+	+		+	5-1	Numbers
B	5	F	AM	+	+		+	4-6	
B	7	F	PM	+	+		+	5-0	
C	6	F	AM	+	+	just perfect	+	5-3	
C	8	F	PM	+	+		+	5-0	
C	9	F	PM	+	+		+	5-1	
C	4	M	AM	+	+		+	5-1	Prepositions
B	5	F	AM	+	+		+	4-6	
B	7	F	PM	+	+		+	5-0	
C	6	F	AM	+	-	and give the box to me	+	5-3	
C	8	F	PM	+	+	just perfect	+	5-0	
C	9	F	PM	+	-	and give the box to me	+	5-1	
C	4	M	AM	+	-	and give the box to me	+	5-1	Directions
B	5	F	AM	+	+		+	4-6	
B	7	F	PM	+	+	I'm sorry-Cindy where is	+	5-0	
C	6	F	AM	+	+		+	5-3	
C	8	F	PM	+	+	marvelous-there on this picture	+	5-0	
C	9	F	PM	+	+		+	5-1	
C	4	M	AM	+	+		+	5-1	Concepts

Child #	Testing Time	Child #	Testing Time
5	5'08"	8	7'23"
7	8'40"	9	6'50"
6	6'25"	4	8'10"

Table 1

Summary of Accuracy of DIAL Administrative Procedures - Communications

Operator	Child #	Sex	Time of Day	Provides Warm Climate	Uses Exact Wording	Word Additions Or Changes	Enters & Exits According to Manual	Child Age	Area Subtests
D	12	M	AM	+	+	say it louder for me	+	4-6	
D	14	F	AM	+	+	good for you	+	4-7	
D	10	M	PM	+	+		+	3-4	
D	13	M	PM	+	+	take finger out of mouth	+	5-1	Articulation
D	12	M	AM	+	+		+	4-6	
D	14	F	AM	+	+	beautiful	+	4-7	
D	10	M	PM	+	+	what did I say	+	3-4	
D	13	M	PM	+	+		+	5-1	Memory
D	12	M	AM	+	+	what did you say	+	4-6	
D	14	F	AM	+	+		+	4-7	
D	10	M	PM	+	+		+	3-4	
D	13	M	PM	+	+		+	5-1	Language
D	12	M	AM	+	-	if (replaces when)	+	4-6	
D	14	F	AM	+	-	if (replaces when)	+	4-7	Problem
D	10	M	PM	+	+		+	[illegible]	
D	13	M	PM	+	+		+	5-1	Solving
D	12	M	AM	+	+		+	4-6	
D	14	F	AM	+	+		+	4-7	Self-
D	10	M	PM	+	+		+	3-4	
D	13	M	PM	+	+		+	5-1	Identification
D	12	M	AM	+	+		+	4-6	
D	14	F	AM	+	+		+	4-7	
D	10	M	PM	+	+		+	3-4	
D	13	M	PM	+	+		+	5-1	Classification
D	12	M	AM	+	+		+	4-6	
D	14	F	AM	+	+		+	4-7	
D	10	M	PM	+	+		+	3-4	
D	13	M	PM	+	+		+	5-1	Story-picture

Child #	Testing Time	Child #	Testing Time
12	6'05"	10	8'00"
14	6'00"	13	5'25"

Reliability of DIAL Scoring

Method. On April 1, 1974 a videotape was made of eight white children being screened with DIAL. In each of the four DIAL areas, one boy and one girl, one of whom would be entering kindergarten in the fall while the other would have another year to wait before entering kindergarten, were filmed. In addition to diversity in sex and age, the participation of some children who were causing concern to either parents or the school was requested.

The videotape was scored by C. Mardell and D. Goldenberg. The results of this scoring are referred to as the standard score. The scoring of each person or operator who participated in the study was then compared to this standard score to determine the percentage of agreement.

Subsequently, videotape was viewed by 16 individuals with little or no previous experience with the DIAL who were given instruction in DIAL scoring.

Findings and Conclusions. The number of possible responses in each area range from 28 in Fine Motor to 71 in Communications with Gross Motor having 38 responses and Concepts having 45 responses. Since each operator scored two children in a given area, the number of responses scored must be doubled. The sixteen operators scored a grand total of 3380 responses, with 251 responses differing from the standard score. Table 1 indicates the percentage of agreement by child and by area.

Table 1

Percent Agreement in
Scoring Between Pairs of Scorers

Area	Child Number	% Agreement
Gross Motor	1	.81
	2	.90
Fine Motor	1	.83
	2	.91
Concepts	1	.95
	2	.99
Communications	1	.91
	2	.97

It can be readily seen that the scoring of the first child in each of the four areas is not as reliable as the scoring of the second child in each area. This may be due to the fact that the first child was the younger child, still a year and a half away from the kindergarten, or due to the anxiety level of the operator and/or the scoring of the videotape itself.

In the Gross Motor area, the most unreliable items were scoring of "throwing", "hopping" and "balance beam". After the manual directions are revised, it will be possible to establish if reliability can be significantly increased on these two items.

In the Fine Motor area, the scoring directions for drawing shapes and letters need revision. In the Concepts area, errors in scoring "front" and "back" may be avoided through a manual clarification. Otherwise, this area had particularly high reliability.

Finally, in the Communications area it is difficult to establish reliability on the "story picture" because operators have the option to use a sentence given by the child on an earlier item. In addition, certain visual stimuli need improvement such as pictures of "tail", "match", "rat" and the actions of the bird, horse, and fish.

In general, we believe that the above results are very encouraging with respect to scoring reliability for the DIAL. With the modifications of the manual suggested in a later section that reliability can be expected to become even greater.

II. PROJECT ACTIVITIES AND FINDINGS

D. RECOMMENDED MODIFICATIONS OF THE DIAL

Modifications of DIAL Manual

A number of minor modifications of the DIAL test manual are described below. These modifications are mainly manual clarification procedures rather than major testing revisions. Recommendations for revisions by the various consultants are listed according to the following code: A-Authors, C-Caldwell, Cl-Clapper, G-Goldman, H-Halverson, J-Johnson, Jo-Jordan, K-King, W-White, Wi-Wick.

Page iii Contents -- add section on Parent Communications, page 7, between High Risk, page 6, and Registration, page 9. (A)

Page 1 Administration -- under General: add the model for Screening and Diagnosis. Team Specifications: delete parenthesis items for Operators and Volunteers. Add bracket and (professional and/or paraprofessional) next to operators (5). (A,K ?) Professional: #3 -- insert: "any of the following areas:" before special education. (A) Coordinator: #5 -- delete: "and among", "and OSPI DIAL staff". Insert: "and other interested parties". (A)

Page 2 DIAL Floor Plan (A)

Paragraph 1: Delete sentence beginning "only playdough . . . area".

Paragraph 2: Delete "Have"; begin "Camera film, bulbs and wastebasket should be available at the photo station."

Paragraph 3: Delete and insert:

"The gross motor station should be placed as seen on Diagram I.

The drum is placed against a wall. A piece of red tape is placed on the floor in front of the drum. Three feet are measured from the taped line and marked with a second red tape line. The child will stand at the tape three feet from the drum and toss the bean-bags to the drum."

Page 3 Diagram I (A, H)

1) Add wall and line above drum and [O] [C] at three foot line.
2) Add [V] or [O] at Photo.
3) Add [C] [O] or [V] above play table.
4) Add [Co] . Delete [C] at registration table.
5) Add key: [O] = Operator, [C] = Child, [V] = Volunteer, [Co] = Coordinator.

Page 5

Information: delete "site number, child number".

Recording of Scores: Paragraph 1, sentence 2 delete "by" and insert "in".

Paragraph 2: delete sentence beginning "For instance... is circled". Continue with sentence beginning "Each operator... each task". Add "If a child self corrects, the best possible score is recorded."

Order: Paragraph 1: add "The child may move to any open station once he or she is comfortable. For a hesitant or extremely quiet child, delay the communication area for the last station."

Page 6 (A) Paragraph 1: add "or at eye level on the nearest wall".

Orientation: Paragraph 1: To sentence beginning "This allows...",

insert "or nickname" between "name" and "and".

Paragraph 2: Darken type of sentence beginning "No matter... progress."

High Risk: Delete first four words and begin sentence "The co-ordinator... screening." Last sentence, last paragraph, delete after "months" and insert "which total under 13 points or falls into a high risk range."

Page 7 (A,Wi) Paragraph 1, last sentence: insert "or with a high risk score in the communication area" between "stations" and "should".

Insert Section -- Parent Communication

It may be advisable to phone or mail the results of the DIAL testing a day or two following the screening. Redial and or diagnostic appointments may be scheduled before the child and parent leave the screening. At no time should the DIAL scores be discussed in the screening area among other adults, neighbors and children. Privacy should be maintained in as best a manner as possible.

Page 9 (A) Preparation: (a) add "/coordinator/volunteer". Directions: delete first two words and insert "The coordinator welcomes parent and child." In the third sentence beginning "The operator will...", delete the first three words and insert this sentence: "The child's name or nickname and chronological age in years and months (i.e., 4-3) are printed onto a name tag from information provided by the parent." Last paragraph, same section,

delete entire last sentence beginning "After the... area."

Paragraph 2: Begin "An operator or volunteer will be seated among the children at the play table to establish a warm relaxed atmosphere." Second sentence, insert "or volunteer" between "operator" and "is". Sentence 3, delete "the operator is cautioned" and insert "caution is taken" between "However" and "not". Sentence 4, insert "or volunteer" between "operator" and "at". Delete sentence 5 beginning "A volunteer... operator." Delete sentence 7 beginning "The volunteer... print." Sentence 8 should begin "The child's full name is printed on the back of the photo." Sentence 9 should begin "This packet moves with the child from station to station."

Page 10 (A) Photographers Instructions: Paragraph 1: add sentence 2, "The photograph should include the child's body."

Paragraph 4: delete the first two sentences entirely and move the remaining sentence beginning "The child. . . tion" to follow the last sentence in paragraph 3. Delete the last sentence on page 10.

Page 11 (H,A,W,C) Gross Motor: In reviewing the integral components of this area, it would seem appropriate to re-locate the identification of body parts (Item #7) to the Concepts area.

There is justification for removal of the growth measures until such time as there are scales developed for prediction purposes.

Preparation: Item (g) -- substitute the word "drum" for "cylinder" throughout the gross motor area and delete (j), (k).

Directions:

1) Throwing: delete (1) (H,A,Wi). Change successive numbers to begin with (1) and continue through (4). Delete the overhand beanbag toss (H,A,Wi).

(1) Operator says "Hi, (child's name), we're going to toss bean bags. Watch this."

(2) Operator tosses, underhand, 3 bean bags in the direction of the drum, standing behind taped line 3 feet away from drum.

(3) Operator picks up tossed bean bags and placed them on floor directly in front of child. If child does nothing, operator says, "Your turn."

(4) After child tosses bean bags, operator collects them. Operator circles hand(s) child threw with and records child's score.

Scoring: Correct possible points to equal 3.

2) Catching: (H,A)

Scoring: add "even if he steps across line." "(possible points = 3)" to follow the word child.

Page 12 3. Jumping (H,A) Scoring: Three points for landing on both feet at the same time, two points for any jump attempt;

4. Hopping (H,A) (2) change six to five. Scoring: change word "each" to "continuous" hops and underscore five per leg. (Possible points = 5 right + 5 left.)

5. Skipping (H,A) (2) change "six" to "three". Scoring: change to read "three points for a step/hop coordinated skip; two points for a slide or gallop; one point for any other attempt;" Add: (Possible points = 3)

6. Standing Still (A) (3) Underscore arms or legs and substitute "30" for "20". Scoring: change to read "Three points for standing still 30 seconds, two points for standing still 20-29 seconds, one point for standing still 10-19 seconds. (Possible points = 3)

7. Body Parts Move body parts to Concepts (H,W,A,Wi)

Page 13

8. Balance Beam (H,W,A,Wi) Redefine title of task to describe activity. (W) Change to Balancing. Delete (1) and (2). Child should not walk the beam in stocking feet (H,Cl,A,Jo). Re-number remaining items in a consecutive manner. Add sentence "Operator can allow child a second trial if there is a faulty start", to follow step 3.

9. Growth Measures (H,W,A,Wi,Jo) Delete total item.

Page 15 Fine Motor: There should be a consistent use of task description in the sub-test heading and text. Each task activity should be reflected in the area sub-test. (W,Q,Jo)

Designs will become Matching

Blocks will become Building

Cutting will remain Cutting

Drawing will become Copying Shapes and Copying Letters

Finger Agility will become Touching Fingers

Hand Clapping will become Clapping Hands

1) Matching (6) Add: "If child responds by pointing, operator takes child's hand and places disc in position. Operator underlines each shape matched correctly. The scoresheet should have each design in order or presentation." (7) Item listed as 6 becomes 7.

Page 16 Scoring: add (Possible points = 3) (A)

4. Copying Shapes

5. Insert Copying Letters (A,W). Items listed as 8-10 under Drawing become Step 1-3 under letters.

Scoring: Move scoring of shapes to directly follow the section of the text and have scoring of letters directly follow Copying Letters text. Insert scoring criteria for both letters and shapes.

Page 17 6. Touching Fingers (A) Scoring: add (Possible points = 3)

7. Clapping Hands (A) Scoring: add (Possible points = 3)

Page 19 Concepts -- Learning Process will become Sorting Blocks (W,A,Jo)

Colors will become Naming Colors

Numbers will become Counting

Prepositions will become Positioning

Following Directions will stay the same

Concepts will become Identifying Concepts

1. Sorting Blocks (A) Step 2. Insert "Then it will be your turn' (A) following the first sentence ending "me".

Page 20

3. Counting (A) Add: step 3 -- operator circles number child counts consecutively up to eight. Delete scoring which follows step 3. Steps 3-8 to become 4-9. (Step 6, replace "give me" with "take". Operator replaces all blocks after each task.) Step 9 -- add: (1, 3 or 5). After step 9, add: "Scoring: One point for each numeral counted consecutively (Possible points = 8)" and "One

point for each block correctly placed (Possible points = 5).

4. Positioning (W,A) Step 4 -- add: The block must be placed in back of or in front of the box according to the child's perspective. Scoring: add (Possible points = 14).

Page 21 Move body parts from Gross Motor (W,A,H,Wi)

7. Identifying Body Parts

Step (1) Operator stands up near table and says to child "Come over here."

Step (2) Operator says "O.K., (child's name), show me your finger." If child does, go to Step 5.

Step (3) - Step (8) follow in a consecutive order as they appear under Gross Motor.

Page 23 Communications (W,A) Sub-test headings will change in headings and throughout the text.

Articulation will become Articulating

Memory will become Remembering

Language will become Naming Nouns and Verbs

Problem Solving will become Coping

Self-Identification will become Naming Self, Age and Sex

Classification will become Classifying Foods

Story Picture will become Telling a Story

Articulating (G,K,Wi,A) Step 3 -- delete and insert "Operator records a 2 for each correct response and a 1 for each partially correct response." Step 8 -- change underlines to circles.

Page 24 Naming Nouns and Verbs Add the list of acceptable responses.

Page 25 Coping (A) Step 4 -- Scoring: add (Possible points = 6) Naming Self, Age and Sex: (A,Wi) Step 9 -- delete beginning "Operator... through scoring". Scoring: add "and one point for correct sex" to follow "age". (Possible points = 4)

Page 26 Classifying Foods (A) Step 4 -- insert the word "different" between "of" and "foods". Add "Brand names are acceptable." Scoring: add (Possible points = 6)
Telling a Story (A) Scoring: one point for each part of speech circled (Possible points = 7)

Other Possible Revisions for the Production of DIAL III

Several major item revisions may be desireable in the development of DIAL III. These future additions or substitutions should follow clinical pilot studies which analyze the effects of added variables. In addition to revisions listed by area, it is important to record the length of time required for the completion of each station. This potentially useful information is not recorded under the present DIAL format.

Gross Motor

Within the conceptual design of the DIAL format, processing of information is recorded as well as the end product. This could be improved in the gross motor station where only the end product is scored. Different methods of recording this process should be investigated. It would be helpful to know the effect of removal of the drum and an accuracy throw rather than a distance throw. Different levels of competence would have to be established for

the total age range. It would be advantageous to investigate alternative methods for stimulating a jump and a hop response. Since the balance beam presents a safety question, studies should analyze the effect of shoes on or off, the beam being sprayed with a non-slip product or the beam being covered with a carpet or cloth surface.

Fine Motor

Combining both gross and fine motor stations may reduce the length of testing time. Since the gross motor area takes the least amount of time, a combination of the two areas may facilitate the screening of large populations. Inclusion of an intermediate block building task such as a six-block tower prior to the six-block pyramid could provide more developmental information. Since the present scoring system for shapes and letters create errors, a procedure for simplification should be investigated. There should be a review of the order and complexity effect of clapping patterns and, possibly, new combinations could be tested keeping in mind the goals of developmental temporal sequencing.

Concepts

The concepts area may require additions of items which reflect pre-classification skills and associations. Notation of scoring of the learning task should allow for a recording of process as well as product. It would seem useful to devote some in depth study to the selection of pictorial representations for the concepts being sampled. The addition of body parts to this area should be studied in relation to High Risk scores being affected by the order movement of this item.

Communications

Since verbal behavior is so vital for school success, this section deserves the closest review. Again, the pictorial representation for the items within this area needs study.

A more complex memory task for digits should be included and alternative sentences for the meaningful memory section could be added.

Since there were many diverse questions concerning the acceptable responses for solving problems, the possibilities of added acceptable answers should be pursued.

Elimination of the Polaroid picture may be a necessary consideration because of the expense of photographic materials. It would be important, however, to study the response of other self-identification methods such as a mirror or simply posing the question without a visual stimulus. The added value of the picture for later staff efficiency, in review of large numbers of children screened in a single day, could off-set this cost priority decision.

The final item, the story picture, requires the greatest effort. Eliciting a story from a stationary picture may not provide enough of a stimulus for the desired oral response. It may be advisable to allow the use of all expressive language offered in the communication area as the corpus of speech and direct the testor to score sentence length and structure from the total sample. Presentations of a story picture should include studies of the effect of perpendicular or flat presentation.

Social Affective Behaviors

This area is in need of much further investigation. It may be inappropriate to attempt to measure these behaviors in the limited time allowed for screening. Also, there should be some collection of the number of DIAL II behaviors recorded on high risk children as opposed to children scoring above the risk prediction. There may be some means in which a child's awareness of others' needs can be measured.

III. CONCLUSIONS AND RECOMMENDATIONS

Summary of Conclusions

The several conclusions stated below are based both on current project findings and on information obtained previously. These conclusions form the bases for the subsequent recommendations concerning the future development and use of the DIAL.

1. The content of the current DIAL is generally consistent with current information, theory, and opinion regarding child development. The only major gap in that content is in the area of social-emotional behavior.

2. Application of the training procedures for DIAL operators devised earlier by C. Mardell and D. Goldenberg is likely to result in sufficiently high reliability of test administration and scoring procedures.

3. The predictive power (validity) of the present battery (DIAL II) very likely is sufficiently high to warrant use of the battery for screening purposes with children four years and older Actual widespread use of the battery with such children probably ought to be delayed until 1975-76 when it will be possible to confirm (or disconfirm) that conclusion and to describe the extent of the relationship between DIAL II performance and subsequent school success more precisely.

4. There is some indication of adequate predictive validity of the DIAL battery for use with children from 2.5 to 4 years of age but that indication is based on extremely incomplete infor-

mation. Thus, additional information is required and widespread use of the DIAL II with such children should be delayed until 1976-77, pending confirming data.

5. Development of a revision of DIAL II is advisable. A number of modifications recommended in earlier reports concerning the DIAL and several others generated during this present contract period are warranted. Such a revision (a DIAL III) can be produced in a relatively short period (i.e., four-six months), but the compilation of norms and data concerning reliability and validity will require an additional four years. However, the effort is very likely to result in a battery clearly superior to the present DIAL and with more adequate data in support of its utility.

6. Revisions of the scoring system for DIAL (recommended earlier by J. Wick) will result in higher predictive validity of DIAL scores than will be obtained with current scoring practices.

7. It probably is feasible and desireable to use a two-stage process for screening in which use of the full DIAL battery constitutes the second stage. The first stage is likely to consist of the use of a teacher questionnaire, or the administration of a subset of items from the DIAL battery itself. It seems likely that such a pre-screening stage can quickly, inexpensively, and reliably identify a substantial proportion of children in the screening age range who will require no additional attention.

8. Long-range planning concerning the DIAL and its use has been inadequate. The DIAL (or any screening procedure) ought to be viewed as one element in a <u>system</u> of identification-treatment.

Such a system is sufficiently important and complex to require strong statewide leadership. The mechanism for such leadership has not existed and apparently still does not exist. Without it the goals reflected in House Bill 323 are unlikely to be achieved.

Recommendations

The data that we have obtained this year along with past information suggest that further development of the DIAL and its eventual widespread use are desireable. We again recommend a plan that will provide data sufficient to begin the widespread use of the present battery within two years, and the production and dissemination of a final version of the battery by 1978. That recommendation is based upon several facts and/or assumptions: 1) The availability of a carefully developed screening instrument is important, 2) No such instrument currently exists, 3) The DIAL instrument shows promise of filling that need with additional revision and study, 4) Even without such revision the DIAL instrument is likely to be useful at least on an interim basis.

The recommendations to be described below emphasize careful data gathering and revision work and activities to ensure widespread and proper use of both the current battery and the recommended revision. Our recommendations are designed to accomplish three goals: 1) to make available a useful screening instrument as soon as possible, which we believe will require two years; 2) to make available a fully validated instrument of higher quality than the present version of the DIAL battery, a task that will require four years; and 3) to ensure that as each of these

instruments becomes available its potential and limitations will be understood and it will be used effectively to benefit the children of Illinois.

Recommendation 1

We recommend that additional predictive validity data be gathered over the next two years concerning the present DIAL battery.

During the present contract period all of the children tested previously with the present DIAL battery were in kindergarten, namely all those children who were below 54 months of age when tested. A judgement regarding school success seems tenuous indeed at that level. Thus, although we have collected data regarding such children we must be extremely cautious in drawing conclusions regarding the predictive validity of the DIAL battery. We can only say that there is some suggestion of predictive power and cannot begin to estimate the magnitude of the relationship. We did obtain criterion data for first-graders, but the children in this case had been tested with an earlier version of the present DIAL battery, and all were 4.5 years or older when tested. These data indicate a substantial relationship between scores on that earlier DIAL battery and school success for children 4.5 years and above when first tested. We suspect that the relationship would be stronger with the present battery, but we cannot be certain of that nor can we estimate the extent of the superiority.

Clearly, if the present DIAL battery is to used widely in Illinois considerable additional information is desireable. That

information can be obtained during the 1974-75 and 1975-76 academic years by further follow-up of the children tested during 1972-73. The follow-up would take a form similar to that used in the follow-up of first-graders under the present contract. The criterion data would consist of achievement test scores supplemented by teacher judgements. Such data would be obtained in 1974-75 for those children 4.5 years or older when administered the DIAL battery and the following year for the remainder of the children. These data would be sufficiently complete so that if they indicated reasonably high validity the DIAL battery could be put into use at that time. There is one significant limitation of these data that would temper conclusions. At the time the DIAL battery was administered it was common practice to provide parents and/or teachers information regarding their children's performance. It is possible, therefore that such "contamination" could alter the course of events so as to either inflate or decrease the relationship found between DIAL scores and subsequent school success. Although we suspect that such effects would be minor, that is an assumption that is untestable and certainly weakens the information regarding predictive validity. Nevertheless, we believe that the follow-up data will provide an adequate basis for a judgement regarding at least the short-term use of the present DIAL battery.

Recommendation 2

We recommend that a revised version of the DIAL battery be produced to replace the present battery to function as the primary preschool screening battery in Illinois.

A number of modifications of the present battery appear desireable. Many of these alterations are minor, consisting of changes in instructions for given items and the like. Some, though, involve the possible elimination of items (e.g. the balance beam) or the development of new items (e.g. the possible construction and testing of new cognitive items as suggested by Burton White). There seems little doubt that a revision of the battery based on information already available would result in a substantially improved screening instrument. To produce that instrument and provide the necessary supporting reliability and validity data require a period of four years. In the first year (1974-75) the anticipated revisions can be made and the instrument administered to a sample of children on whom validity data will be gathered. Data concerning reliability of administration and scoring also can be obtained during that year. During the third year (1976-77) criterion data can be obtained on the older children in that sample, who would be in Grade 1 at that point. During 1977-78 such data can be collected for the younger children in the sample, and second-grade performance can be assessed for the older portion of the sample. At the end of 1977-78 sufficient data would be available for widespread use of the instrument as a replacement for the present DIAL battery.

Recommendation 3

We recommend that the OSPI support sufficient information and training activities to ensure the widespread proper use of the DIAL batteries as they become available.

Recommendations 1 and 2 are designed to make available preschool screening procedures that will be useful to schools and beneficial to children. However, the mere availability of such procedures does not ensure such outcomes. The history of testing, particularly intelligence testing, is ripe with examples of the misuse of tests and test data. To ensure that the DIAL batteries are used widely and with maximum effectiveness requires, we believe, that procedures be established for the dissemination of information and materials and for the training of key school district personnel, i.e., the school official responsible for preschool screening within the district.

This task is a complex one, and we are not able at this point to describe the required procedures in full detail. However, several points will be essential in their development. First, such procedures must be both effective in reaching appropriate school personnel and feasible in terms of the resources of the State. Thus, it seems unreasonable to expect the OSPI to assume the major responsibility for all future training activities concerning the administration and interpretation of the DIAL batteries. Rather, procedures should be developed by which school districts can effectively assume such responsibilities with some limited outside guidance. One approach to consider here is

to offer workshops to be attended by one or two key individuals from any given school system. Such individuals then would be in a position to take effective leadership regarding the use of the DIAL battery within their respective systems. The workshops would center on preparing participants to train others to administer the battery and to make correct use of the resulting data within a system of screening, diagnosis, and intervention.

Second, the procedures should be timed to coincide with the introduction of the current DIAL battery into general use in Illinois. The major efforts, then, would occur during the final three of the four years proposed earlier. The first two of those three years would concentrate on the use of the current DIAL battery and the final year on the introduction of the revised battery. We believe that the switchover from the current to the final battery can be handled without extensive retraining of personnel. What will be required is a highly effective program for dissemination of the revised form and the resulting changes in administration procedures and interpretation. We believe that procedures can be developed such that by the end of 1977-78 no further training activities sponsored by the OSPI will be required.

Recommendation 4

The OSPI should make plans for the publication of the DIAL instrument. We suggest making contact with private publishing companies within the next year to begin movement in that direction.

Recommendation 5

We recommend the development of a two-stage screening system in which the administration of the full DIAL would constitute the second stage.

What we are suggesting here is that adequate screening probably does not require the administration of the complete DIAL battery to all children in the screening age range. We suspect that very brief and inexpensive "prescreening" procedures can be used to identify a relatively large proportion of preschool children that will almost certainly encounter no special difficulties in school. It is very likely that the administration of a selected small number of items from the DIAL battery can serve that purpose. Other possibilities include, for children enrolled in preschool programs, use of a teacher questionnaire.

Recommendation 6

A committee or task force should be established to plan and oversee the development of a system of identification and intervention in which screening is one important element.

This perhaps is our most important recommendation since its implementation is necessary to the development of both the DIAL and the system in which the DIAL would be used effectively. The erratic nature of the development of the DIAL to date stems largely from the absence of the planning and general supervision we are emphasizing here. We believe that this can be corrected by establishing a strong core organization within the OSPI that

undertakes those responsibilities. This body should be headed by an individual who has policy planning and implementation responsibilities within OSPI. It should include others lower in the hierarchy whose responsibilities are at the implementation and evaluation level. It also should include, perhaps in an advisory or consulting capacity, individuals outside OSPI who can provide specialized input appropriate to the needs of the group at any given point in time. For example, during the test development stage advice from specialists in measurement and test development is critical. Later, or perhaps concurrently, advice from specialists in systems development and analysis would be extremely useful in dealing with certain decisions.

This body should meet frequently and regularly, generating policy recommendations (e.g., for legislation), developing requests for proposals when appropriate, reviewing progress of contracted activities, and so forth. It should outline a long-range plan and develop plans for the phases by which the long-term objectives can be reached.

CONCURRENT DIAGNOSTIC EVALUATION OF DIAL II CHILDREN

The purpose of this study was to establish and record the individual diagnostic performances of a discrete number of children who had participated in DIAL II screening. The diagnostic results for those children predicted as potential high risk candidates and their achievement levels were assembled and tallied for two school districts who participated in the 1973 screening.

Since the prediction of high risk preschool children presents a myriad of problems in the selection of appropriate measures for the establishment of normal developmental achievement, it was important to consider an alternative of a concurrent diagnostic comparison between those children identified by DIAL and accompanying results gained from individual psychological testing.

Northbrook

One of the sub-sites participating in this project was Northbrook District #28. As part of North Suburban Special Education District Site 20, they were given training materials and evaluation. District #28 randomly screened 322 children. According to the DIAL II cut-off points, there were 14 (5%) high risk, 30 (11%) who needed re-dialing, and 278 children (84%) who seemed to be progressing adequately.

During the fall of 1973, Project DIAL staff contacted District #28 to continue the validation study of those children screened in the previous spring. Each child's original score

sheet was checked for scoring of computational errors. The 14 high risk children had been recommended for further study to verify the high risk prediction. This was done prior to staffing conferences with the parents of the high risk children. Each parent of a potential high risk child was given the test information results. During this conference, further evaluation plans were presented and the parent was given several options. The results of the diagnostic evaluation are depicted in Table 1.

Further Validation of High Risk Prediction

Within DIAL prediction lies the possibilities of correctly or incorrectly identifying potential school problems and preventing school failure.

There were 14 children identified as high risk. One child (#1) had no achievement information available and two children had moved out of the district (#10, #14). Of the remaining 11 children, seven were entering kindergarten and four were pre-kindergarten.

Of the four prekindergarten children, one child (#2) was evaluated at Northwestern University Learning Disabilities Clinic in November, 1972, and diagnosed as a delayed language case. This child's parent refused district special education placement and the child received special private therapy. In this case, DIAL II identification of high risk was accurately substantiated by clinical diagnosis.

Child #8 was evaluated by the Northbrook Instructional Examiner, Mrs. Joellen Mack, on April 22, 1974. His mother had requested a formal speech evaluation in December, 1972. Testing

RESULTS OF DIAGNOSTIC EVALUATION OF

NORTHBROOK DISTRICT #28

DIAL-HIGH RISK SUBJECTS

CHILD #	SCHOOL GRADE	SERVICES	METRO TEST%
1	K	No testing prior to K	No results
2	pre-K	Northwestern U. clinic	
3	K	Special education-LD	65%
4	K	Special education-LD	65%
5	K	Hearing loss-reg. program	67%
6	K	Parental rejection	73%
7	K	Bilingual language difficulty	67%
8	pre-K		
9	K	Parental rejection	30%
10	K	Child moved	86%
11	pre-K	Parental rejection	
12	pre-K	Family therapy-Joselyn Clinic	
13	K	Parental rejection	69%
14	pre-K	Child moved	

TABLE 1

was attempted at that time but not completed. It was recommended that he remain in the nursery school setting in which he was participating and return to the district DIAL screening in April of 1973. During this evaluation, the child fell below the cut-off points in all DIAL areas. The child participated in the district speech therapy program beginning October, 1973. He had received both small group and individual speech assistance. His recent diagnostic testing showed a WPPSI Verbal IQ score of 85 and a Performance score of 100 for a Full Scale score of 91. It was felt that the speech therapy the child was receiving was good but the high risk performance prompted a more comprehensive plan for future educational programs. Following the diagnostic assessment, it was recommended that this child be considered for acceptance into the Special Education Cooperative (NSSED) child development room. If he does not qualify for that placement, he will receive additional language stimulation within the program guidelines of District 28.

Child #11 was identified as a potential high risk but the parents refused further evaluation.

Child #12 would not have been identified as high risk by DIAL II scores. He scored below cut-off points in gross and fine motor skills which would normally call for specific recommendations to be made to the parents. However, he was viewed as a potential high risk because of the district's knowledge of the family's previous emotional problems and private treatment.

Of the nine kindergarten children screened, Child #3 and Child #4 (identical twins) were evaluated on June 25, 1973. During

the DIAL testing, the parents contributed information regarding a history of ear infections and requested a thorough psychological evaluation prior to kindergarten.

Child #3 received a Verbal IQ score of 86 and a Performance IQ score of 100 for a Full Scale IQ score of 92 on the WPPSI. The Peabody score showed low receptive language. All of the auditory abilities on the ITPA were 6 months to a year below age level. The child did not understand directions and did not receive a score on the Detroit tests. Her performances on the diagnostic battery resulted in a recommendation for learning disability placement. Her DIAL high risk score was substantiated through diagnostic evaluation.

Child #4 received a Verbal IQ score of 84 and a Performance IQ score of 108 for a Full Scale score of 95 on the WPPSI. She was approximately a year below her chronological age level on the ITPA sub-score dealing with Verbal Abilities and 6 months below her chronological age level on the Manual Expressive Abilities. Her Peabody Vocabulary score was 86 with a mental age of 4-0, chronological age was 5 years, 2 months. The instructional examiner recommended a learning diability placement based upon the diagnostic confirmation of the DIAL II prediction of high risk.

Child #5 was screened on April 12, 1973. His chronological age was four years, four months and fifteen days. The low cut-off point was for the communication area, alone. Auxillary information added to his test protocol confirmed no special testing referral at this time. Child #5's mother was recently deceased and a S.L.I.D.E.S. hearing evaluation detected a slight hearing loss. Recommendations

were made to review his file at a later date.

Child #6 was a child who scored below the DIAL II cut-off point in two areas on April 11, 1973. Her chronological age was four years, five months. After retesting in the Communication and Fine Motor areas, she still remained below the cut-off points, therefore becoming a potential high risk. The parents rejected further evaluation. They described the child as shy and timid. Once entered in kindergarten (Fall 1973), further evaluation was requested by the kindergarten teacher. The results of that evaluation were not available for inclusion in this report.

Child #7 was screened on April 12, 1973. Her chronological age was five years, three months. At that time she scored below the cut-off points in two areas, communication and concepts. Further information from an interview with the parents described a bilingual learning situation within the home. The parents were Ukrainian and the child had difficulty learning English when Ukrainian was spoken at home. No special testing was recommended at that time, however placement within a regular nursery program was suggested.

Child #9 was screened with the DIAL II battery on April 3, 1973. At that time he was four years and eight months. Based upon the DIAL II scores Child 9 was below cut-off points in three areas, Gross Motor, Fine Motor and Concepts. Discussion with the parents resulted in a rejection of further diagnostic evaluation prior to kindergarten. He entered kindergarten and was referred

for testing by his kindergarten teacher due to difficulty with following directions.

Diagnostic evaluation was accomplished on December 10, 1973. No visual or hearing difficulties were reported.

Child 9 was given the WPPSI and achieved a score of 92 on verbal sub tests, falling in the lower end of the average range, and 111 on the performance sub tests, placing him in the bright range. The nineteen-point difference between verbal and performance scores indicated a possible learning problem.

Child 9's performance on the Detroit Tests of Learning Aptitudes was a year or more below age level in five of the sub tests.

ITPA performance was one year or more below age level in five sub tests.

The Wepman Auditory Discrimination scores were inadequate for his age level. Bender designs were quite disorganized and the Frostig test score showed two areas of performance with one year or more below age level. The instructional examiner recommended assistance from the learning disabilities teacher.

Diagnostic testing in this instance. confirmed the DIAL II high risk prediction.

Child #13 was screened on April 11, 1973. At that time he was four years and nine months. His low scores were in the gross motor and communication areas. The mother rejected further evaluation because of the child's shyness.

DIAL II Validation -- False Negatives

Child #188 was screened with the DIAL battery on April 9, 1973. At that time he was five years and one month. Based upon the

research cut-off points, Child #188 did not fall below the criterion number in any of the four areas.

The parent's request for a psychologist's evaluation of this child on May 11, 1973 at chronological age five years, two months resulted in the following data:

Stanford Binet	CA-5yr.,2mo.;MA-5yr.,5mo.;IQ-105
Cooperative Pre-School Inventory	Raw Score 42-63%
Peabody Picture Vocabulary	MA-5yr.,1mo.; IQ-99
Test of Visual Motor Integration	VMI age-5yr.
Vineland Social Maturity Scale	Social age-6yr.,10mo.

The psychologist did not recommend a special Child Development placement for this child but rather recommended a structured kindergarten and consultation services for guidance of social behavior patterns. This child displayed rather intense and high strung behaviors of "testing limits".

It is rather difficult to interpret DIAL predictive validity based upon current information. Obviously this child is not achieving to potential as evaluated by the Metropolitan Readiness score. If a criterion of school success includes factors of school adjustment and social behaviors then the crucial factor of integration variables creates a muddled view of this child's true school performance.

In this case, DIAL II score did not accurately predict the social behavior difficulties mentioned by the psychologist. The predictive validity study of DIAL I included first grade achievement scores and teacher evaluation. This child scored in the twenty-ninth percentile on the Metropolitan Readiness test. Hypothetical suppositions as to the reasons for such performance

would be inappropriate with the limited available validity data.

Child 188 is only one of the two false negatives found in the Northbrook study. There was no diagnostic information on Child #293, who was the second false negative DIAL child. This child was screened on April 3, 1973, and was five years, zero months at that time. His DIAL II scores did not fall below the research cut-off point for his age and sex, yet he scored in the thirty-fifth percentile on the Metropolitan Readiness test.

Wilmette

District 39, Wilmette Public Schools, was also a sub-site member of Site #20 for DIAL II testing in 1973. Wilmette was a non-random site which screened children thought to have developmental lags. The publicity distributed in the community encouraged parents to bring children to a testing site for vision, hearing and DIAL II evaluation. Parents who had concerns about their children's developmental status were offered consultative services.

In the Spring of 1974, Wilmette was contacted to supply further validation of DIAL II prediction for children screened during the Spring of 1973.

District 39 held their early childhood registration on a "walk-in" basis and all forms were completed for vision and hearing appointments as well as DIAL appointments. Forty-nine children were seen for DIAL II assessment and District 39 initiated a Child Development Program to service sixteen children identified as needing special services. This number increased to nineteen with the inclusion of three mid-year screened children.

Placement in the special services child development program was determined by follow-up diagnostic evaluations of those children

falling into the high risk range as determined by any low DIAL area score (Table 2). On occasion, difficult decisions were resolved through the use of additional clinical judgements supplied by the DIAL II screening staff and supplemental information.

The Child Development program began in September, 1973, and continued through June, 1974. The staff consisted of a Child Development classroom teacher who saw the children four hours per week and a resource team (nurse, social worker, speech and language therapist, psychologist, learning disability teacher, classroom teacher and Assistant Superintendent for Pupil Services) which provided monthly program and progress assessments as well as professional support.

Almost daily contacts were made with parents and two formal parent conferences were scheduled. Following the conference, parents were given a copy of the Progress Report.

In May, 1974, all children enrolled in the program were re-screened with the DIAL II instrument (Table 3). Based upon test performance, year long observation, resource team judgments and parent opinions the folloiwng figures represent the future class placements of these children (Table 4).

The Assistant Superintendent for Pupil Services aad discussed the efficiency and accuracy of DIAL II screening with the district's regular program nursery staff at the mid-year. The staff reported that there were no regular nursery program children in need of special services and that the screening program did a thorough job.

WILMETTE DIAL II TEST/RETEST

1973 Test	Gross Motor	Fine Motor	Concept	Communications	Cut-Off	C-A
Child #1	10	4	7	6	5	2-10
2	0	1	4	0	--	2- 3
3	-- (9)	6	-- (19)	-- (9)	6	3- 1
4	--	--	5	--	6	2- 9
5	2	--	5	1	4	2- 5
6	8	9	10	8	5	2-11
7						
8						
9						
10	14	5	14	10	7	3-10
11						
12	--	--	--	--	6	2-11
13	8	8	7	10	9	3- 7
14	12 (11)	11 (10)	-- (14)	-- (15)	12	3- 9
15	-- (0)	-- (0)	-- (0)	-- (0)	5	2- 7
16	6 (0)	1 (4)	6 (7)	7 (7)	5	2-11

TABLE 2

WILMETTE DIAL II TEST/RETEST

1974 Test	Gross Motor	Fine Motor	Concept	Communications	Cut-Off	C-A
Child #1	15	15	15	13	12	3-12
2	10	7	7	6	7	3- 5
3	18	17	20	18	12	4- 2
4	--	--	--	--	12	3- 9
5	9	13	17	15	7	3- 5
6	15	16	20	14	12	3-11
7	12	9	15	16	15	
8						
9						
10	16	19	20	20	14	4- 3
11						
12	16	18	13	16	13	3-11
13	15	16	14	16	14	4- 7
14	18	17	21	17	16	4- 9
15	9	13	17	13	10	3- 8
16	14	14	15	16	12	4- 0

TABLE 3

Future Class Placement of

Child Development Class (1973-4)

Children entering first grade	2
Children entering kindergarten	3
Children entering nursery school	2
Children leaving District 39	2
Children returning to the program	10
Total	19

TABLE 4

The follow-up of this site provides further documentation of the broad range of "screening philosophy" that exists in the field today.

District 39 had made a commitment to developing preventative measures for children displaying developmental lags. Rather than a concentration on the identifying of severe lags, this district selected children for a diagnostic evaluation who displayed low scores in any DIAL area.

Results

The results of this study offer strong support for continued refinement of an early developmental screening device. Based upon these limited but encouraging studies, the DIAL instrument shows promise in providing a screening tool which can identify and predict those children exhibiting a developmental delay.

It is difficult to evaluate predictive validity of DIAL II based upon the fact that the Wilmette sample was given special assistance and the Northbrook population provided small numbers of kindergarten children who had participated in DIAL screening. Predictive studies present a methodological paradox. If early screening and identification were successful and an early childhood program provided those tasks or experiences which assisted a child toward more positive school efforts, the child would no longer be high risk. He or she would be a competent school achiever. Therefore, establishing the predictive validity of the screening device that succeeded in accurate identification would, optimally, be a negatively precise one.

Since no control group was established at the time of the Wilmette Child Development group, it is also difficult to separate the factors which may have contributed to a child's higher retest DIAL II score. The factor of maturation should influence success on DIAL II scores since all DIAL II items were initially selected for their developmental correlation. It would then be expected that a 4-year-old child would perform better on each task merely by virtue of age.

Diagnostic evaluation of potential high risk children can provide specific remediation avenues for early amelioration of severe learning problems. Reviewing the collection of information concerning the fourteen children termed "high risk" within the District #28, DIAL screening provides further support for second level diagnosis of a potential learning problem prior to special services placement.

A screening test, at best, should over-identify rather than under-identify those potential learning problems within the larger population. Accuracy of identification is of the utmost importance; however, a screening system is expected to be a gross measure and the diagnostic battery a refining system.

The criterion for school success established for DIAL II predicaion actually rested upon a child's first grade achievement performance and a teacher evaluation. Any prediction validation based upon kindergarten readiness factors alone is questionable. Also, the DIAL II sample had feed-back in the form of scores and prediction validity should be established without the contamination of a known or expected performance.

If diagnostic evaluation can be used as an added criterion for accuracy in prediction of potential learning problems then of the four Northbrook "high risk" children receiving diagnostic evaluation, all were confirmed to have specific speech, language or developmental problems which were reflected in a developmental performance delay.

It does confuse the issue to refer to readiness scores of children who have received some special assistance and then attempt to project the quantity or quality of their achievement performance had no help been provided. Also, there is the age-old question of later maturation. If left alone, would the DIAL II high risk child achieve success in an academic climate without special assistance?

Surely this small study of DIAL II children only heightens the concern for careful assessment and evaluation of pre-kindergarten children. At no time should screening identify or label children. Developmental performance screening should be viewed as a positive ongoing process, reviewed with parents and with specific recommendations made for further educational assistance. Screening to identify problems without further guidance and services to the children in need is a disservice to both the children and the community.

SCHOOL DISTRICT
65
Cook County
EVANSTON, ILLINOIS

Ida B. Lalor, Director
Research and Evaluation

MEMORANDUM

TO: Dr. James Hall
FROM: Ida B. Lalor

RE: Summative Evaluation Report on the DIAL project

As outlined under State of Illinois Contract No. L171, two evaluators, external to the project, were to determine the extent to which the project staff had met the goals of further developing and refining the DIAL instrument. As one of the evaluators I will comment on the project as follows:

Efforts undertaken to provide for

(a) construct and
(b) predictive validity of DIAL
(c) standardization and simplification of procedures for administering the instrument so as to increase the reliability of the observations made,
(d) improvement of the questionnaires for the collection of observations from parents and teachers, and staff evaluation of the effectiveness of the questionnaires in identifying high risk students.

(a) Construct Validity

Extreme care was taken by project personnel to insure a rigorous test of the theoretical basis for DIAL.

i) An evaluation of the instrument was made by experts in the areas of the behaviors being measured and

ii) the instrument was compared to other approaches and practices used in diagnostic centers across the country.

Before these activities were undertaken, criteria for the selection of experts and sites to be visited were developed.

Copies of the selection criteria, the dates, locations and outcomes of the consultations and visits were provided the evaluators. It is my judgment that the methods employed were comprehensive, appropriate and highly productive.

(b) Predictive Validity

Despite the many problems associated with retrieving a previously tested population of students, the project staff was able to collect current achievement data for 249 children who had been screened on DIAL during 1971-72 and 1972-73

through school records or by administering the Iowa Tests of Basic Skills. That instrument was employed in those instances in which achievement scores were not available. In addition to achievement scores, teacher responses to a Classroom Performance Assessment Scale were collected.

While there was some delay experienced primarily due to the need to convert scores to a common scale and achieve a consistent data format, the analyses are in process and findings will be included in the final report. The staff is on the threshold of identifying those items in DIAL which differentiate between children who are likely to succeed in school and those who are not unless some formal intervention is provided. Since the population sample was stratified by sex, race, socio-economic status and urban-rural location, the data can be further interpreted in relation to these factors.

(c) Inter and Intra-rater Reliabilities

As has been characteristic of the previous activities described, the study on inter-rater reliabilities was carefully planned. Videotapes of children in test situations were used as data sources to determine the consistency of scoring of students' performance by a given evaluator using DIAL criteria at two points in time. The expense of using this medium placed constraints on the initial design of the study. Thus far, the only analysis has been to determine the per cent agreement between scores and rescores given by the same raters; these are very high. The staff has decided to make an additional video tape to make possible other combinations of students and examiners and to derive a coefficient of correlation between score-rescore by raters.

(d) Parent and Teacher Questionnaires

In my opinion, the one aspect of the project that is not progressing at a rate comparable to the others, is the refinement of the parent and teacher questionnaires. Different kinds of information have been collected from teachers and from parents using several instruments. The relationships between these instruments and DIAL should be determined. The feasibility and validity of substituting a questionnaire for the screening procedure is of considerable significance and deserves special attention. There may well be far-reaching implications for teacher and parent education.

The Future of DIAL

The project director has laid out reasonable alternative courses of action concerning the future of DIAL, all of which start with the premise that DIAL has the potential for meeting an observed and expressed need in the field of early childhood education. As a researcher, I agree with the judgment that additional predictive validity data be collected and that a likely outcome of such studies would be a further revision of DIAL. As a practitioner, however, I am somewhat less cautious and, based on the construct validity of the instrument, feel secure in suggesting that the DIAL instrument, modified to incorporate improved procedures for administering and scoring, be used while the projected further studies are in process.

As the State advises local education agencies of the need to provide special education services for children between the ages of three and five years, it can advise local education agencies about or furnish the screening procedure that will enable agencies to determine if a diagnostic assessment is indicated or not.

At the same time, the State ought to mandate that a screening instrument, however valid, may not be used for placing children in special education programs. A single administration of DIAL can tell us if a child's development seems to be following a typical pattern, at a given point in time, or not. If the system is used as designed, the consequences of incorrectly identifying a child as a high risk would be to take that child through a diagnostic assessment procedure unnecessarily. The economic impact of such a misidentification would be alleviated by coordinating the local education agencies' screening procedures with community diagnostic clinics. Conversely, the assumption is that, given the care with which the instrument has been constructed, few high risk children will fail to be identified. However, without making use of the instrument, more high risk children would be missed. The value of using DIAL at this stage of development seems evident.

In sum, the evaluation of the methods and products used and developed by this program leads me to suggest that the office of the Superintendent of Public Instruction provide school districts with an integrated set of policies and suggested procedures, including DIAL, that would facilitate the implementation of House Bill 323 of the 77th General Assembly.

Ida B. La Cour

Ida B. Lalor, Director
Research & Evaluation

SUMMATIVE EVALUATION REPORT

OF THE PROJECT

"FURTHER DEVELOPMENT AND REFINEMENT OF DIAL"

Report Prepared by:

Dr. Harold J. McGrady
Director of Special Education
Mesa Public Schools
39 S. Hibbert St.
Mesa, Arizona 85202

Report Submitted to:

Office of the Superintendent of Public Instruction
State of Illinois

October, 1974

This summative evaluation is concerned with the activities of the DIAL Project since November, 1973. However, it seems appropriate to place the current year's activities in full perspective of the two years' work which preceded this past year. The 1973-74 effort appears to have been a natural evolution of the work that preceded. Thus, the procedures and results of this recent work should be evaluated in light of, and in relation to the preliminary activities of 1971-1973.

Continuity

In reviewing three years of effort, one important factor emerges, namely that the primary objective has remained the same: that a screening procedure be designed to identify "high-risk" pre-kindergarten children, i.e., those who are most likely to demonstrate learning disabilities at school age. It would appear that the three-year progress has been remarkable in moving toward that final goal. A review of that progress is assessed below.

A second critical factor has been the continuity of key personnel involved with the project. Dr. Mardell and Mrs. Goldenberg, who together initiated the project in 1971 have continued throughout the entire life of the project. Furthermore, Dr. Wick has completed nearly three years as a project staff person. These members of the staff have been unrelenting in their efforts to bring in many consultants, evaluators, and the like from a wide variety of disciplines, always carefully selecting such personnel with stated criteria. The record also shows that they have heeded the advice generated from such consultations. Planning for subsequent years has always taken into consideration the available counsel.

Long-term Planning

What will happen at the completion of this three-year project? What is the long-term commitment to the goal stated as the primary objective? The need for a screening system to pick out potential learning disabilities (of all types) is a given. The State of Illinois has mandated services for handicatpped children in the pre-school years down to age three. It necessarily follows that children who are in need of pre-school special education should be identified. If this is to be done, the initial DIAL efforts should be continued and expanded. This takes commitment of financial and personnel resources. If such a commitment is to occur, the leadership must come from sources high enough within OSPI to ensure that the commitment can be honored.

A great deal has been learned about screening by the DIAL staff. That effort should not be dissipated. During the past three years the DIAL Project has continued on a year-to-year basis, with no evident master plan for future commitment. It would appear that OSPI should engage in an effort now to decide at the highest administrative and decision-making level just what will be done with DIAL. My review of the project efforts is favorable toward the continuation of an effort that has already yielded many results for the expenditure of relatively little money (when compared to other conparable research and development activities.

Need for a System for Screening

A commitment for the development of a pre-school screening instrument by OSPI has some advantage, namely that it could be plugged into a state-wide system for screening, identification, diagnosis, and special education.

This implies, of course, that OSPI will develop a master plan for such a system, and that the screening instrument could serve a specific function within that total system and could be utilized Statewide in the implementation of that system. Not only would such an effort serve the need to identify all pre-school children who are in need of special education, but it would represent a leadership function in this area of endeavor, probably unparalleled in this country.

If a total screening and identification system is planned it must take into account a) need for other types of screening in education, medicine, and psychology; b) the systems to be used for identification and diagnosis after pre-school screening has occurred; and c) what delivery systems will be available for treatment or educational intervention. (It does no good to invent a screening and diagnostic system that classifies children into discrete categories and then places them all in the same treatment.)

Of course, the self-evident problems of financing, operationalizing, and administering such a system must be considered. Not the least of all problems to be considered in planning such a system is how we can insure that all children who need to be treated are entered into the screening and identification system.

Early Stages of DIAL

But these factors are looking more to the future of DIAL and the probability that the original primary goal can be met. Let us review briefly the accomplishments of the project before the current year.

The first year was devoted primarily to searching the literature to find ready-made techniques for meeting the objective. It was determined

that a new instrument would need to be developed. This was accomplished, and the instrument was field-tested. From the field tests certain items were selected as being potentially useful. It was recognized, however, that there was a need to verify the usefulness of such items through further study, primarily to assess the reliability, validity, and predictability of the newly-developed instrument.

In reviewing the reports submitted by evaluators at the completion of the first year, this evaluator was impressed with how seriously the project personnel have taken the suggestions made by various evaluators. Virtually all such suggestions have been put into effect during the ensuing two years. This reinforces the need for constant formative evaluation, and supports the positive effects that can result when such advice itself is evaluated and implemented whenever feasible.

How does DIAL Fit into a Total Scheme of Screening?

But, heeding the advice of previous evaluators has not solved all problems (nor should that be expected). A problem that has persisted has been an adequate statement, definition, or operational definition of what types of "learning disabilities" are to be detected, or how specifically they should be detected by the screening instrument. Let me state it in a different manner.

In any screening system we must ask, "What are we screening for?" That is, "What problems are we trying to detect?" Two types of screening might be envisaged: (a) gradual grading, or (b) differental sorting. each type of system poses different questions.

In (a), the "grading" system, we keep asking the same question over and over again. ("Does the child have a learning disability?"), only we use a finer and finer instrument at each stage of screening and identification. As in the case of sifting rocks out of sand, we can use a large

gauge screen if we only want to eliminate (or identify) larger rocks. Everything that is smaller will pass through and we will consider as acceptable. Only the large rocks will be considered "defective", because they could not pass our "screen". Subsequently, if we need to select smaller and smaller rocks and pebbles, we can simply apply a more stringent screen, one with smaller holes than the previous one(s).

Applying the gradual grading type of system to learning disabilities it simply would mean that we ask the same question over and over ("Does the child have a learning disability?") If we apply a very coarse, or gross screen initially, we will identify only those who have severe learning disabilities, then we gradually apply a finer and finer screen (or evaluation) until even subtle disabilities will be selected.

(The reverse grading could also occur.. That is, we could ask the reverse question: "Does the child not have a learning disability?" Then we could select a very fine grade screen, so that only those about whom no question exists would be allowed to pass. The process of grading would then continue in reverse of before, going from fine to gross grades, till only the severe disorders were left.)

In either grading system, one could arbitrarily decide upon a level of severity as the cut-off and use only one sifting operations. By this method, in one step a child would either be included or excluded from further consideration.

In (b) the differential sorting system, we do not keep asking the same question over and over again. Rather we systematically ask a series of different questions, each designed to detect a different quality or characteristic. Ultimately this would be differential diagnosis, and is analogous to a qualitative analysis in chemistry. The questions are sequenial to determine whether the object has characteristic A; char-

acteristic B, etc. The sum total of answers provides an exact designation of what the substance is. In learning disabilities screening, identification and diagnosis, we simply ask a series of differentiating questions, e.g., does he have hearing loss? Is he low in concept ability? Can he count? etc. The types of questions and the discreteness with which we answer them depend on what we are looking for. Again the sum total of all of our answers tells us what we have.

Screening for children with learning disabilities could incorporate a combination of both the gradual grading and the differential sorting systems.

I do not believe that the DIAL project has taken this type of a systems analysis approach to the problem. That is, they have not clearly defined the decision making process of which the DIAL instrument is intended to be a part. If this question is addressed we can then determine the validity of DIAL in a more comprehensive manner. That is, we can test how valid DIAL is as a device in the total decision-making process.

It would appear that DIAL attempts both a grading and a differential sorting. The sampling of various types of behaviors is a type of differential sorting, but the use of cut-off points is a type of grading. It is difficult to determine which of these aspects is more validly accomplished, but would appear to be a favorable characteristic of the instrument that each type is incorporated.

This raises another question that appears to be recurring, that is, whether prediction or detection is desired. Again, let me use an analogy. In medical screening or diagnosis we could aim to detect a disease at its earliest possible stage to keep it from spreading, as is the case in cancer. For example, a lump is direct evidence of an abnormal growth. In contrast, we use predictions to prevent developement of other diseases,

most notably heart disease. It is well known that stress, smoking and overweight are conditions which increase the probability of heart disease. By decreasing these conditions we decrease the probability of heart problems. In this case our treatment, or intervention is based on prediction, not early detection. Obviously either method could be used for either disease: heart disease or cancer. That is heart disease could be detected in early stages before major damage occured or cancer could be prevented by heeding predicting factors, such as the known information that heavy smoking leads to higher probabilities of contacting lung cancer.

In screening for learning disabilities we must determine whether we are going to engage in prediction or early detection. The decision will determine what types of measures we use, when we will apply our test, and what we expect to do with the results.

If our goal is to predict then we must state what we want to predict and how early such prediction should occur. Only then can we determine the predictors, which often will be behaviors different from the target behavior which we are concerned about. For example, we might find that not having an oral vocabulary of such-and-such a size at a certain age would be highly predictive of reading failure at a given grade in school. If we could determine that we could intervene by attempting to increase vocabulary after noting its level through our "screening device". If we were certain of this predictive relationship we would only need to devise a reliable, valid measure for assessment of vocabulary and our decision about intervention would be dictated by our screening and diagnostic results. The point is that our screening technique would emerge from our determination that oral vocabulary level is predictively related to eventual reading level.

97

In this instance we measure one type of behavior to predict and intervene in our concern for a different behavior. In contrast when we set early detection as our goal, we will be measuring the target behavior,

but measuring it at a convenient time so that if the behavior is below par we can intervene directly. Thus, in the case of reading as a traget behavior, we assess reading skills as they are taught to determine progress and the need or lack of need for intervention. This means setting up periodic testing of the target behaviors.

Again, it appears that DIAL utilizes both of these approaches in its operation. And this has been so undefined that there have been some confusions in the "validity studies". In some cases, it appears that prediction has been the goal, because attempts are being made to correlate behaviors at one age (eg gross and fine motor ability) with different target behaviors at another age (e.g. reading achievement during 1st grade). In other instances, the early detection has been the goal, because there has been a stated concern for determining the child's "ability to abstract" at pre-kindergarten, which would emerge at school age as a continuing problem in ability to abstract. Perhaps a key ingredient in DIAL is that it purports to do both prediction and early detection. The full validity for either of these is still unproven, although the results in so far are very promising for both prediction and detection.

Perhaps both processes need to be a part of the total system, but the use of DIAL should be identified and described in relation to the total decision-making process. A specific example of where this has led to confusion is in relation to the developement of "criteria for school success". In some instances this has been interpreted as an independently derived index of whether the child has a problem. The concept was apparently raised by one evaluator of the first year of the project (see p. 97 of the 1st year report). He indicated that it might be fruitful to analyze what skills are needed to succeed in school; then determine the extent to which DIAL scores are related to those skills. This was not the way in which the DIAL staff interpreted and

analyzed "criteria for school success". Instead they asked consultants to derive lists of behaviors which denoted school success and then merely asked whether their DIAL items were measuring the same or similar behaviors at an earlier stage. Thus, the DIAL analysis was one which concerned itself with face validity of their items; whereas the previous evaluator appeared to be suggesting that an analysis should be done to test the predictive validity of their items against measures of skills known to be necessary to succeed in school. The latter has not been accomplished.

A similar confusion exists in the use of the teacher questionnaire in the current year. The teachers rated the children on behaviors intended to be distinctive from academic school success, which in turn was noted through the achievement test scores. Thus, the correlations of DIAL scores to achievement scores indicate the ability of DIAL to predict academic success, but the correlation of DIAL scores with teacher ratings indicates the ability to predict other-than-academic behavior (usually social-emotional-behavioral skills) from DIAL. There was a feeling from our discussions with DIAL staff that there was an expectation that all three measures (i.e. DIAL, academic achievement, and teacher ratings) should be highly intercorrelated. Understandably, they did not, because the teachers rated non-academic behavior. Thus DIAL correlated more highly with academic achievement and less with social behaviors, ala the teacher ratings. Consequently teacher ratings and academic achievement were not highly correlated. Thus DIAL emerges as a predictor of academic school success per se in contrast to being a predictor of good school behavior. This is primarily due to the original purpose of DIAL. If good school behavior is to be predicted different items would need to be incorporated into DIAL.

The above has been a discussion of general trends and issues that

that have appeared throughout the history of DIAL. Let us now look specifically at other issues and activities from the current year.

Evaluation of 1973-74 activities

Because the objective and activities stated in the proposal were modified to some degree, we will not attempt to evaluate each of these. Rather, we will confine our statements to actual activities accomplished. Let us state however, that the decision making process which led from proposed to actual activities was a logical one and was monitored throughout by the project formative evaluator.

Validity

Our prominent group of activities was aimed at determining the validity of the nature of the instrument, especially whether the "right" abilities were being measured. Experts were queried and literature was viewed to determine whether the items on the test were appropriate for early childhood screening and whether they were representative of behavior for the target behavior to be measured. The DIAL staff made a thorough literature search, and they were unrelenting in their efforts to identify and consult with "experts" who could add to their ideation regarding the test battery makeup. The results suggest that a) the construct validity is logically sound, i.e., a comparison of DIAL items with other test items show similarities in what they intended to measure; and b) the face validity is high, in that most experts agreed that the items selected would be predictive of school success.

Thus, after careful selections of consultants, thorough review of literature, and selected observations of the screening procedures used by others, the results indicated that the basic concept of DIAL, its constructs, and purposes are logically valid and supportable. Their findings should provide impetus for future development, instrumentation, and validation.

Further validation data are forthcoming, but at the time of our evaluation, some data were not yet analyzed. Thus, no comment will be made regarding the criterion-related validity study.

Test simplification

The studies designed to simplify test administration seemed to be progressing adequately, although this effort may be a wasted one at this time. It would seem more appropriate to agree on a final version in regard to validity and predictability before tightening up a "tester-proof" version. The previous work by Wick has established the generally good reliability of the items. This again is a plus in favor of continuing the development of DIAL.

Particularly promising is Wick's attempt to provide weighted scores that add to the usefulness as a screening instrument. This approach should be pursued strongly. If carried out it should allow for the development of a longitudinally predictive scale of considerable value. I stress again the need to continue, rather than drop the effort of three years.

Conclusions

In total it appears that the development and validation of DIAL has proceeded logically and that the activities during the past year have been entirely appropriate for the stage of evolution in which we currently find the overall project. The concept and need for early identification of children with learning problems is a reality. The constructs which would appear to be best suited for measurement by such a scale seem to be readily agreed upon by serious workers in this field and fields of related study. Suitable items have been field-tested, validated and proven to be reliable. Refinement of technique, item selection, and score weighting have each evolved to a relatively sophisticated level. In fact, in many ways, it might be said that the

current DIAL instrument is as finely-tuned an instrument as is available for its avowed purpose anywhere in this country. It only remains to continue that refinement, verify the validity over longer periods of time in given children, and clarify the interpetation of results so that DIAL can be used within the proper perspective of a total screening system.

In future development the following questions should be addressed:

a) What types of children is DIAL intended to detect and/or predict.

b) Where in the total screening, identification and/or diagnosis decision-making system does DIAL appear?

c) How well does DIAL serve the above-mentioned purposes?

d) Is DIAL superior to other alternatives?

Past accomplishments suggest positive and useful answers to there questions if only continuity and commitment continue!